

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 12:46:11 ; Search time 7196 Seconds
(without alignments)
10938.038 Million cell updates/sec

Title: US-10-018-902-3

Perfect score: 1924

Sequence: 1 ccacgcctccgcacatcaataa.....aaaaaaaaaaaaaaaaaaag 1924

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pi.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_on.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1924	100.0	1924	6	AX081073	AX081073 Sequence
2	1216.2	64.3	1703	8	AB011416	AB011416 Oryza sat
3	1216.4	63.2	1938	8	HVGTRNARI	X92403 H.vulgare m
4	1215.6	63.2	1997	8	HVHEMARI	X86101 H.vulgare m
5	1214.2	63.1	1864	8	D88382	D88382 Hordeum vul
6	1196.4	62.2	1778	6	AX081077	AX081077 Sequence
7	1107	57.5	1480	6	AX081089	AX081089 Sequence
8	1100.4	57.2	1822	8	D88383	D88383 Hordeum vul
9	1090.6	56.7	1703	8	HVHEMARI	X86102 H.vulgare m
10	987	51.3	130058	8	AC078840	AC078840 Oryza sat
11	987	51.3	300029	8	AE017109	AE017109 Oryza sat
12	710.4	36.9	2055	6	AX081081	AX081081 Sequence
13	706.2	36.7	2072	8	D50407	D50407 Cucurbit m
14	679.6	35.3	1632	6	AX081085	AX081085 Sequence
15	679.6	35.3	1663	8	AY096600	AY096600 Arabidops
16	679.6	35.3	2111	8	AY072223	AY072223 Arabidops
17	619.8	32.2	1593	6	AX651934	AX651934 Sequence
18	602	31.3	1901	8	D67088	D67088 Cucurbit m
19	589.2	30.6	1983	6	AX081085	AX081085 Sequence
20	588.6	29.0	3663	8	AF105221	AF105221 Glycine m
21	541.2	28.1	86014	8	F19C14	U03774 Arabidops
22	534.6	27.8	2879	8	ATU03774	U03774 Arabidops
23	509.6	26.5	1979	8	AF294753	AF294753 Hordeum v
24	492.2	25.1	118803	8	AC124214	AC124214 Medicago
25	475.6	24.7	1047	8	AB075573	AB075573 Glyptost
26	472.4	24.6	1047	8	AB075574	AB075574 Taxodium
27	467	24.3	3683	8	AB075575	AB075575 Metasequo
28	467	24.3	3683	8	ATU27118	U27118 Arabidops
29	467	24.3	146505	8	F21M12	AC000132 Sequence
30	466	24.2	1047	8	AB075580	AB075580 Juniperus
31	464.4	24.1	1047	8	AB075576	AB075576 Sequoia s
32	464.4	24.1	1047	8	AB075582	AB075582 Cryptomer
33	461.2	24.0	1047	8	AB075577	AB075577 Chamaecyp
34	461.2	24.0	1047	8	AB075579	AB075579 Thuja sta
35	461.2	24.0	1047	8	AB075581	AB075581 Chamaecyp
36	458	23.8	1047	8	AB075578	AB075578 Thujaopsi
37	444	23.1	3179	8	AF305613	AF305613 Chlamydom
38	380.6	19.8	84872	8	AC006593	AC006593 Arabidops
39	329	17.1	344	8	AF496730	AF496730 Zea mays
40	329	17.1	344	8	AF496731	AF496731 Zea mays
41	329	17.1	344	8	AF496732	AF496732 Zea mays
42	329	17.1	344	8	AF496733	AF496733 Zea mays
43	329	17.1	344	8	AF496734	AF496734 Zea mays
44	329	17.1	344	8	AF496735	AF496735 Zea mays
45	329	17.1	344	8	AF496736	AF496736 Zea mays

ALIGNMENTS

RESULT 1
AX081073
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AX081073
Sequence 3 from Patent WO0109304.
AX081073
AX081073.1 GI:13170000

Zea mays
Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.

1
Cahoon, R.E., Gutteridge, S., Harvell, L.T., Rafalski, J.A., Tao, Y. and
Weng, Z.

FEATURES	Polynucleotides encoding aminolevulinic acid biosynthetic enzymes									
	Patent: WO 0109304-A 3 08-FEB-2001;									
	E.I. DU PONT DE NEMOURS AND COMPANY (US)									
source	Location/Qualifiers									
	1..1924									
	/organism="Zea mays" /mol_type="genomic DNA" /db_xref="taxon:4577"									
BASE COUNT	504	a	465	c	551	g	404	t		
ORIGIN										
Query Match	100.0%; Score 1924; DB 6; Length 1924;									
Best Local Similarity	100.0%; Pred. No. 0;									
Matches 1924; Conservative	0; Mismatches 0; Indels 0; Gaps 0									
Qy	1	CCACGCGTCCGCATCAATAAAGAGAGAGCTTGGGAAGTTGCCAAGGCCTCCGATTTCGCTA	60							
Db	1	CCACGCGTCCGCATCAATAAAGAGAGAGCTTGGGAAGTTGCCAAGGCCTCCGATTTCGCTA	60							
Qy	61	ATGCGACGATAATGCGGACCAACAGCGTCAGCGACCACCGCCGCGCAGCAGCGCGCACCA	120							
Db	61	ATGCGACGATAATGCGGACCAACAGCGTCAGCGACCACCGCCGCGCAGCAGCGCGCACCA	120							
Qy	121	CCGCCAAGCGCGGGGTCTGTCGCGCCCTCTGCCAGAGGGTGGCCGGCGCGCAGCGC	180							
Db	121	CCGCCAAGCGCGGGGTCTGTCGCGCCCTCTGCCAGAGGGTGGCCGGCGCGCAGCGC	180							
Qy	181	GGCGCTCCGGGTGTGTCGGTGCAGCGCGCGCGGTGGAGGCCCGAGCGCAGCGCGCGTGG	240							
Db	181	GGCGCTCCGGGTGTGTCGGTGCAGCGCGCGCGGTGGAGGCCCGAGCGCAGCGCGCGTGG	240							
Qy	241	CAAAAGGCTGCCAGCGTCGCCCGCCCTCGAGCAGTTCAGAGATATCCCGCCGACCGGTACATGA	300							
Db	241	CAAAAGGCTGCCAGCGTCGCCCGCCCTCGAGCAGTTCAGAGATATCCCGCCGACCGGTACATGA	300							
Qy	301	AGGAAAGAGTACCATAGCTGTGATAGCGCTCAGTGTATACACAGACACAGTGGAGATGC	360							
Db	301	AGGAAAGAGTACCATAGCTGTGATAGCGCTCAGTGTATACACAGACACAGTGGAGATGC	360							
Qy	361	GTGAAAAACTTGCTTTTGCAGAGAACTGTGGCCCGGTGCTATTCAAGAACTCACTAGCC	420							
Db	361	GTGAAAAACTTGCTTTTGCAGAGAACTGTGGCCCGGTGCTATTCAAGAACTCACTAGCC	420							
Qy	421	TGAACCATATTGAAGAGCGTCTGTTCTTAGTACCTGTAAATAGAAATTTATGTGG	480							
Db	421	TGAACCATATTGAAGAGCGTCTGTTCTTAGTACCTGTAAATAGAAATTTATGTGG	480							
Qy	481	TGGCGCTATCATGGAAACCGTGTATCAGAGAAGTAGTGGACTTGAATGTCTGAAGAAAGTG	540							
Db	481	TGGCGCTATCATGGAAACCGTGTATCAGAGAAGTAGTGGACTTGAATGTCTGAAGAAAGTG	540							
Qy	541	GTATTCCCGCTTCCGAGCTTAGGGAGCACTGTTTCATCTTGCAGAGCAGTGAATGCCAC	600							
Db	541	GTATTCCCGCTTCCGAGCTTAGGGAGCACTGTTTCATCTTGCAGAGCAGTGAATGCCAC	600							
Qy	601	GCCATCTGTTGAGGTGTCAGCTGGCTTGAATCTTTGGTCTCTGGTGAAGGCAAAATCC	660							
Db	601	GCCATCTGTTGAGGTGTCAGCTGGCTTGAATCTTTGGTCTCTGGTGAAGGCAAAATCC	660							
Qy	661	TTGCTCAGGTTAAACAAGTTGTGAGGAGTGGACAGACAGTGGAGGCTTGGGAAAGAACA	720							
Db	661	TTGCTCAGGTTAAACAAGTTGTGAGGAGTGGACAGACAGTGGAGGCTTGGGAAAGAACA	720							
Qy	721	TGGATAGGATGTCAGGATGCATCACTGTGGAAAGCGTGTCCGACGAGACCAACA	780							
Db	721	TGGATAGGATGTCAGGATGCATCACTGTGGAAAGCGTGTCCGACGAGACCAACA	780							
Qy	781	TATCATCTGGTGCTGTTTCTGTGTCAGTTTCAGCGCGGTTGAACTGGCCCTGATGAAGCTTC	840							
Db	781	TATCATCTGGTGCTGTTTCTGTGTCAGTTTCAGCGCGGTTGAACTGGCCCTGATGAAGCTTC	840							
Qy	841	CGAAGTCTGAAGCACTGTCACTAGGATGCTTCTGATTGGTGTGGTAAAAATGGAAAGC	900							

RESULT 2	
AB011416	
LOCUS	
DEFINITION	Oryza sativa (indica cultivar-group) mRNA for glutamyl-tRNA reductase, complete cds.
ACCESSION	AB011416.1
VERSION	GI:2920319
KEYWORDS	glutamyl-tRNA reductase;
SOURCE	Oryza sativa (indica cultivar-group)
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE	1. Nakayashiki, T. and Inokuchi, H.
AUTHORS	Nakayashiki, T. and Inokuchi, H.
TITLE	Nucleotide sequence of a cDNA clone encoding glutamyl-tRNA reductase from rice (accession No. AB011416) (PCR 98-080)
JOURNAL	Plant Physiol. 117, 332 (1998)
REFERENCE	2 (bases 1 to 1703)
AUTHORS	Nakayashiki, T.
TITLE	Direct Submission
JOURNAL	Submitted (23-FEB-1998) Toru Nakayashiki, Graduate School of Science, Kyoto University, Department of Biophysics, Sakyo-ku, Kitashirakawa, Oiwakecho, Kyoto 606-8502, Japan (E-mail:nakkun@molbio.biophys.kyoto-u.ac.jp, Tel:075-753-4201)
FEATURES	Location/Qualifiers
source	1..1703
	/organism="Oryza sativa (indica cultivar-group)"
	/mol_type="mRNA"
	/cultivar="IR36"
	/db_xref="taxon:39946"
	/dev_stage="5-day-old"
	72..1682
	/codon_start=1
	/product="glutamyl-tRNA reductase"
	/protein_id="BAA25003.1"
	/db_xref="GI:2920320"
	/translation="MMASTTSATAGGAFAAATRRVVGGRRPCARVAAGRRRRSGV VVRCDAGVQAQAVAKAAASVALEQFKISADRYMKERSIAVIGISVHTAPVEMR EKLAABELPRAISELSLNHIEEAAVLSTCNMEIYVVALSNRGLREVVDWMSKK SGIPASELREHLFDSDATDHLFVPSAGLDSLVLGEIEILAQVKQVVRSGNSGGL GKNIDRMFKDAITAGKVRCKTVISSGAVSVSSAAVELAKMLPKSECLSEMLLIGA GKMGLVVKHIIAGCKKVVNRSVERVDALREEMKDIEIVRPLTEWYEAAREADV VFTSPASSTKFTKEHALPALSDAMGVRFLFVDISVPRNVSACSVSGHARVTVND DLKEVVEANKDRLKAMEAQTIITQELKRFEMRDSLETPTTIKKRSYADRIASE LEKCLQKIGEDALTKKTRSEELSTGIWNKLLHGLPQLHRCDSGSDSLDTLELNMH ALNRFSLDTEKATIEQKIKAKVEKSN"
BASE COUNT	416 a 390 c 538 g 359 t
ORIGIN	
Query Match	64.3%; Score 1236.2; DB 8; Length 1703;
Best Local Similarity	86.2%; Pred. No. 1.8e-202;
Matches 1406; Conservative	0; Mismatches 213; Indels 12; Gaps 3;
Qy	68 GATATGCGGACACAGCGTACGACACACAGCGCGCCGAGCAGCGCCACACCGCCAA 127
Db	
Qy	71 GATGATGCGGACACAGCGTACGACACCGCGCGCGCGCGCGCGCGCGCGGAG 130
Db	
Qy	128 GCGCGCGGCG-GTGCTGCTGGCGCTCTGCCA-----GAGGGTGGCGCGCGCGCGCGG 181
Db	
Qy	131 GCGGCGGCTGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGAG 190
Db	
Qy	182 GCGTCCCGGGTGGTGGCGGAGCGCGCGCGCGCGTGG-----AGGCCAGGCGCAGGC 235
Db	
Qy	191 GCGGTCCGCGGTGGTGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCAGGC 250
Db	
Qy	236 CGTGGCAAGAGCTGCACGCTCGCGCGCTCGAGCAGTTCAAGATATCCGCGCGCGCGTA 295
Db	
Qy	251 GGTGCCAAGCGCGCAGCGTGGCGCGCTCGAGCAGTTCAAGATCTCCGCGCGCGCGTA 310
Db	
Qy	296 CATGAAGAAAGGAGTACCATAGCTGTGATAGGCTCTAGTGTACACACAGCACCGAGTGA 355
Db	

311	CATGAAGAAAGGAGTACCATAGCGGTAAATAGCGCTCAGTGTACACACTGCACCACTGGA	370
Qy	356 GATCGGTGAAAAAATTGCTGTTTGCAGAGAACTGTGGCCCCGTGCTATTCAAGAACTCAC	415
Db		
371	GATCGGTGAGAAACTTGTGTTTGCAGAGAACTATATGGCCCCGTGCTATCTCAGAACTCAC	430
Qy	416 TAGCCGTGAACCATATTGAAGAGGCTGCTGTTCTTAGTACCTGTAAATAGAAATTTTA	475
Db		
431	CAGTCTGAATCATATTGAAGAGGCTGCTGTTCTTAGTACCTGTAAATAGAAATCTTA	490
Qy	476 TGTGCTGGCGCTTATCATGGAAACCGTGTATCAGAGAAAGTAGTGGACTGATGTCGAAGAA	535
Db		
491	TGTGCTAGCTTATCTGCGAACCCTGGACTTAGAGAGTGGTAGACTGATGTCAAGAA	550
Qy	536 AAGTGGTATTCCTCCGCTTCGAGCTTAGGGAGCACCTGTTTCATCTTGCAGAGCAGTGATGC	595
Db		
551	AAGTGGAAATCCCTGCTTCTGAGCTCAGGGAGCATCTATTTCATGTTGCGTGACAGTGATGC	610
Qy	596 CACAGCCCATCTGTTTGAAGGTGTACGCTGGCCCTTCACTCTTTGGTCTCGGTGAAGGACA	655
Db		
611	CACAGCCCATCTGTTTGAAGGTATCTGCTGGGCTTCACTCTTTGGTCTTTGAGAGGCGA	670
Qy	656 AATCCTTGCTCAGGTTAAACAAAGTTGTGAGGAGTGGACAGAACAGTGGAGGCTTGGGAAA	715
Db		
671	AATCCTTGCTCAAGTTAAACAAAGTTGTGAGAGTGGGAAAACAGTGGAGGCTTGGGAAA	730
Qy	716 GAACATCATAGGATGTTTCAAGGATGCAATCACTGCTGGAAAGCGTGTCCGAGGAGAC	775
Db		
731	GAACATCATAGGATGTTTCAAGGATGCAATCACTGCTGGAAAGCGTGTCCGCTGCGAGAC	790
Qy	776 CAACATATCATCTGCTGCTGCTTCTGTCAGTTCAGCGGGGGTTGAATGGCCCTCATGAA	835
Db		
791	TAAATATCATCAGGTGCTGCTCTCTGTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG	850
Qy	836 GCTTCCGAAGTCTGAAGCACTGTCTGCTAGGATGCTCTCTGATGCTGCTGCTGCTGCTGCTG	895
Db		
851	GCTTCCGAAGTCTGAAGCACTGTCTGCTAGGATGCTCTCTGATGCTGCTGCTGCTGCTGCTG	910
Qy	896 AAGCTAGTGTATCAACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	955
Db		
911	AAAGTTGCTGCTTAAACATTTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	970
Qy	956 CTCGTTGAAAGGTTGGATGCTATTCGTGAGGAGATGAAAGATATAGAGATCGTGTACAG	1015
Db		
971	TTTCAAGTGAAGGTTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1030
Qy	1016 GCCTCTCTCAGACATGATCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1075
Db		
1031	GCCTCTTACAGAGATGATGAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1090
Qy	1076 ATCTGAAACTTTCATTTTGCAGAAAGAAACAGCAGAGAGCACTCCCTGCTGCTGCTGCTG	1135
Db		
1091	ATCTGAAACTTTCATTTTGCAGAAAGAAACAGCAGAGAGCACTCCCTGCTGCTGCTGCTG	1150
Qy	1136 TATGGAGGTTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1195
Db		
1151	TATGGAGGTTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1210
Qy	1196 GTCTGAAAGTTCGCGCTGCAAGAGTGTCAATGTGCAAGCTTTGAAAGAGTGTGCGAAGC	1255
Db		
1211	GTCTGAAAGTTCGCGCTGCAAGAGTGTCAATGTGCAAGCTTTGAAAGAGTGTGCGAAGC	1270
Qy	1256 CAAACAGAGAGCACTGCTGAGAAAGCAATGAGAGCGCAGCAATCATCATCCAGAGAACT	1315
Db		
1271	CAAACAGAGAGCACTGCTGAGAAAGCAATGAGAGCGCAGCAATCATCATCCAGAGAACT	1330
Qy	1316 GAGAGGTTTCAAGGATGAGAGGACTCGTGGAGACCGTTCCGACCATCAAGAGAGCTGAG	1375
Db		
1331	GAAACGTTTCAAGGATGAGAGGACTCGTGGAGACCGTTCCGACCATCAAGAGAGCTGAG	1390
Qy	1376 GTCTGACGCGACAGGATCAGGCGCTCGAGCTCGAGAGTGCCTGCGAGAGTAGGTGA	1435
Db		

Db 1391 GTCCTACGCCGACAGGATCAGGGCTTCGGAGCTTGAGAAGTGCCTCCAGAAATCGGCGA 1450
QY 1436 GGACGCCCTTACCAAGAGATGAGGAGACCCATCGAGAGCTGAGACCGGATCGTTAA 1495
Db 1451 AGACGCCCTTACCAAGAGACGAGAGATCCATCGAGAGCTCAGCACCGGATCGTAA 1510
QY 1496 CAAGCTCTCCATGGCCCGCTCGACGACCTGAGGTGCGACGCGACGAGCGGACCCCT 1555
Db 1511 CAAGCTCTCCACGGCCCATTCGACGACCTGAGATGTGAGGAGCGACGCGGACCCCT 1570
QY 1556 TGACGAGACCTCGAAGACATGACGCCCTCAACCGGATGTTTCAGCTTCGACATGAGAA 1615
Db 1571 CGATGAGACGCTGGAGAACATGACGCCCTCAACGAGATGTTTCAGCTTCGACACCGAGAA 1630
QY 1616 GCGGATCATCGACGAGAGATCAAGGCCAAGGTGGGAGAGACACAAACTGAGGCCAGGA 1675
Db 1631 GCGGATCATGAGCAGAGATCAAGGCCAAGGTGGGAGAGATCCAGAACTGAGATTGAAG 1690
QY 1676 AGCAATTTTTC 1686
Db 1691 AGAGATTTTTC 1701

RESULT 3
HVGTNRAR1
LOCUS HVGTNRAR1 1938 bp mRNA linear PLN 08-NOV-1996
DEFINITION H.vulgare mRNA for glutamyl-tRNA reductase, 1st isoform.
ACCESSION X92403
VERSION X92403.1 GI:1041424
KEYWORDS glutamyl-tRNA reductase; hema 1 gene; isoform 1.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
1
Bougrì,O. and Grimm,B.
Members of a low-copy number gene family encoding glutamyl-tRNA
reductase are differentially expressed in barley
Plant J. 9 (6), 867-878 (1996)
96267576
8696365
2 (bases 1 to 1938)
Bougrì,O.
Direct Submission
Submitted (16-OCT-1995) O. Bougrì, Institute of Plant Genetics &
Crop, Plant Research, Corrensstr. 3, 06466 Gatersleben, FRG
Location/Qualifiers
1. .1938
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/sub_species="klages"
/db_xref="taxon:112509"
/chromosome="5"
/clone="bhai"
/clone_lib="lamba gtl1"
/dev_stage="12 d old seedlings"
1. .1938
/gene="hema1"
1. .1938
/evidence="experimental"
61. 1644
/gene="hema1"
/note="isoform 1"
/codon_start=1
/product="glutamyl-tRNA reductase"
/protein_id="CAA63140.1"
/db_xref="GI:1041425"
/db_xref="SWISS-PROT:Q42843"
/translation="MAGTSATAAGAFAPAAKARGPAAACPWLVAAAGRRRSGVVRCD
AGGDAQAKASITIALSQFKISADRYMKSSIAVIGLSVHTAPVNRKLAVALAEEL
WPRATISLTHIBEAVALSTCNRMETIYVVALSNRGIREVVDVMSKKGIPASELR

EHLPMLRSDATRHLPFVSAGLDSLVLEGQILAQVQVVRNGQNSGGIGKNIDRMFK
DLITAGKVRCEINISAGVSSAAVELAMKPKSECLSMALLIGAKMGKLVVK
HLIAKCKKVVVRNRSVERDAIEEMKDIIVYRPLTEMYEAAADADVFTSTASES
LLFTKEHAELVLPPIISLAMGVRLFVDISVPRNVGACLSEVHARVYVNDLKEVVEAN
KEDRVKAMBAQITITOLKRFPEARWDSLETPTIKLRSYADRIRASLEKLOKIG
EDNLNKKRRRSIEELSTGIVNKLHLGPIQLHLCRDSRSLDETLENHMLNRMFSLD
TEKAVLEQKIKAKVEKTS"
61. .201
/gene="hema1"
/note="transit-peptide site"
202. .1641
/gene="hema1"
/product="unnamed"
BASE COUNT 494 a 447 c 559 g 438 t
ORIGIN
Query Match 63.2%; Score 1216.4; DB 8; Length 1938;
Best Local Similarity 83.1%; Pred. No. 4.6e-199;
Matches 1426; Conservative 0; Mismatches 276; Indels 14; Gaps 3;
QY 53 TTTTCGCTAATGCGACGATAATGGCGAACCAACGACGTGACGACCAACCGCGCGCAGCAGC 112
Db 42 TTTTGTGCTGCTCTGATGATGGCGGAGCGACGTGAG---CCACGGCAGCGCGCGCGC 98
QY 113 CGCCACCAACGCCCAAGCGCGGGGTGCTGCTGGGCCCTCTGCCAGAGGGTGGCGCGCG 172
Db 99 ATTGCGCCCGCGCAAGCGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 158
QY 173 CGCAGGCGCGCGCTCCGGGTGTCGGGTGCGACCGCGCGCGCGCGCGCGCGCGCGCGCA 232
Db 159 CGCGCGGAGCGGTCCGGCGTCTGCGGTGCGACCGCGCGCGCGGG-----ATGCAC 209
QY 233 GGCCTGCGAAAGGTGCGACGCTCGCGCGCTCGAGCAGTTCAAGATATCGCGCGACCG 292
Db 210 GCGCGCTGCAAGCGCGCGCGCATCACCGCGCTCGAGCAGTTCAAGATCTCGCGACCG 269
QY 293 GTACATGAAGAAAGAGTACCATAGCTGTGATAGCGCTCAGGTACACACAGCAGCCAGT 352
Db 270 GTACATGAAGAAAGAAAGTAGCATCGCTGTATATAGGCGCTCAGGTACACACAGCAGC 329
QY 353 GGAGATGCTGAAAAAATTGCTGTGTCAGAGGAACCTGTGGCGCCCTGCTATTCAAGAACT 412
Db 330 GGAGATGCGGAAAAAATTGCTGTGTCAGAGGAACCTATGGCGCGCTGCTATTTCAGAACT 389
QY 413 CACTAGCCTGAACCATATGAAAGGCTGCTGTTCTTAGTACCTGTAAATAGAAATGAAAT 472
Db 390 CACCACTGTAATCATATTGAAGAGGCTGCTGTTCTTAGTACCTGCAACAGAAATGAAAT 449
QY 473 TTATGCTGCGCGCTATCATGGAACCGTGTATCAGAGGAGTGTGGACTGATGTCGAA 532
Db 450 ATATGCTGCTTTTATCGTGAACCGTGTATTAGAGAGTAGTAGACTGGATGTCAA 509
QY 533 GAAAGTGTATTTCGGCTTCGGAGTTCAGGAGACCTGTTTCATCTTCGGAAGCAGTGA 592
Db 510 GAAAGTGTGAATCCCTGCTTCGAGCTAAGGGAGCATCTCTTATGTTGCGTGACAGTGA 569
QY 593 TGCCACACGCCATCTGTTTGAAGGTGTCAGCTGGCGCTGCACTCTTTGGTTCTCGGTAAAG 652
Db 570 TGCCACACGCCATCTGTTTGAAGGTATCCGCTGGGCTTCACTCTTTGGTTCTTGGAGAGG 629
QY 653 ACAATCTCTCTCAGGTAAACAAGTTGTGAGGAGTGCAGACAGTGTGGAGGCTTGGG 712
Db 630 ACAATCTCTCTCAAGTCAAAACAAGTTGTGAGAAATGGGCAAAACAGTGGAGGCTTGGG 689
QY 713 AAAGAACATCGATAGGATGTTCAAGGATGCAATCACTCTGGAAGCGTTCGCGCAGCGA 772
Db 690 AAAGAACATCGATAGGATGTTCAAGGATGCAATCACTCTGGAAGCGTTCGCGCTGTA 749
QY 773 GACCAACATATCATCTGCTGCTGTTTCTGTCAGTTTCAGCGCGGTGTAACTGGCCCTGAT 832
Db 750 GACCAACATATCATCTGCTGCTGTTCTGTCAGTTTCAGCAGCAGTTGAATTTGGCCATGAT 809
QY 833 GAAGCTCCGAAGTCTGAGACACTGTACAGTAGGATGCTTCTGATTTGGTGTGTTGTAAT 892

QY 190 GGGTGGTGGCGTGGAGCGCCGCGTGGAGGCCCGAGCGCGCGTGGCAAGGCTG 249
Db 451 GCGTGGTGGCGTGGAGCGCCGCGCGG-----ATGCAAGCGCGCGTCCAAAGCGG 501
QY 250 CCAGCGTGGCGCGCTCGAGCAGTTCAAGATATCCGCCGACCGGTGACATGAAGAAAGGA 309
Db 502 CCAGCATCACCGCGCTCGAGCAGTTCAAGATCTCCGACAGCGGTACATGAAGAAAGA 561
QY 310 GTACCATAGCTGTGATAGGCTCTAGTGTACACAGCAACAGTGGAGATGGTGAANAAC 369
Db 562 GTAGCATCGCTGTAAATAGGCTCTAGTGTACACAGCAACAGTGGAGATGGCGGAAAC 621
QY 370 TTGCTGTTGACAGAGAACTGTGGCCCGTGTCTATTCAAGAACTCACTAGCCCTGAACCAT 429
Db 522 TTGCTGTTGACAGAGAACTGTGGCCCGTGTCTATTCAAGAACTCACTAGCCCTGAACCAT 681
QY 430 TTGAAGAGGCTGTCTTTAGTACCTGTGTAATAGAAATGAAATTTATGTGGTGGCTAT 489
Db 682 TTGAAGAGGCTGTCTTTAGTACCTGTGTAATAGAAATGAAATTTATGTGGTGGCTAT 741
QY 490 CATGCAACCGTGTATCAGAAAGTGTGGACTGGATGTCAAGAAAGTGGTATTCGG 549
Db 742 CGTGGAAACCGTGTATTAGAAAGTAGTAGTGGATGTCAAGAAAGTGGATTCGG 801
QY 550 CTTCGAGCTTAGGAGCACCTGTTTCATCTTTCGGAAGCAGTGTGCCACACGCCATCTGT 609
Db 802 CTTCGAGCTTAGGAGCACCTGTTTCATCTTTCGGAAGCAGTGTGCCACACGCCATCTGT 861
QY 610 TTGAGGTGTGAGTGGCCCTCACTCTTTGTTCTCGGTGAAGCAAAATCTTCTCAGG 669
Db 862 TTGAGGTATCGCTGGGCTTGACTCTTTGTTCTTGGAGAGGACAAATCTTCTCAG 921
QY 670 TTAACAAAGTTGTGAGGAGTGGACAGACAGTGGAGGCTTGGAAAGAACATCATAGGA 729
Db 922 TCAACAAAGTTGTGAGAAATGGGCAAAACAGTGGAGGCTTGGAAAGAACATCATAGGA 981
QY 730 TGTTCAAGGATGCAATCACTCTCGAAAGCGTGTCCGACGAGACCAACATATCATCTG 789
Db 982 TGTTCAAGGATGCAATCACTCTCGAAAGCGTGTCCGACGAGACCAACATATCATCTG 1041
QY 790 GTGCTGTTCTGTGAGTTCAGCGGCGTGTAACTTGCCCTCATGAAGTTCGGAAGTCTG 849
Db 1042 GTGCTGTTCTGTGAGTTCAGCGGCGTGTAACTTGCCCTCATGAAGTTCGGAAGTCTG 1101
QY 850 AAGCACTGTGAGTGTGCTTCTGATGTTGCTGTGTAATGGGAAGTGTGATCA 909
Db 1102 AATGTTGTGAGTGTGCTTCTGATGTTGCTGTGTAATGGGAAGTGTGATCA 1161
QY 910 AACATCTGTTGCCAAAGGATGCAAGAGTGTGTTGTTGCAACCGCTCCGTGGAAAGG 969
Db 1162 AACATTTGATTGCCAAAGGATGCAAGAGTGTGTTGTTGCAACCGTTCAGTGGAAAGG 1221
QY 970 TGGATGCTATTGTTGAGAGATGAAGATATAGAGATCGTGTACAGGCTCTCTCAGACA 1029
Db 1222 TGSATGCAATTCGCAAGAGATGAAGATATCGAGATGTGTACAGGCTCTCTACAGAGA 1281
QY 1030 TGTATCAAGCTGTGCTGAAGCTGTATGTTGTTTCCAGCAGCCGCACTCAAACTTCAT 1089
Db 1282 TGTATGAAGCGCTGCTGATGTTGTTGTTTTCAGCAGCCGCACTCAAACTTCAT 1341
QY 1090 TGTTTGGCAAAAGAACACGACAGGCACTCCCGCTGCTCTGTATCTATGAGGAGTGTTC 1149
Db 1342 TATTTCAGAAAGGATGCGAGGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1401
QY 1150 GCCTGTTTGTGCAATATCTCTCCCGAGGATGTACAGGCAATGTGTGTCTGAAGTGGCG 1209
Db 1402 GCCTTTTGTGATATCTCTCTCCAGGAAATGTGCGTGTCTCTCTCTCTCTCTCTCTCT 1461
QY 1210 CTGACAGAGTGTACAATGTGACCACTTTGAAGAGTGTGGAGGCAACAGGAGGACC 1269
Db 1462 ATGACCGGTATACAATGTTGACCACTTTGAAGAGTGTGGAGGCAACAGGAGGACC 1521
QY 1270 GGCTCAGGAAAGCAATGGAGCGCAGACAATCATCACCGAAGAACTGAGACGGTTCGAGG 1329

Db 1522 GTGTCCGAAAGCAATGGAGGCCCAAAACATCATACCCAGAACTGAACGGTTCGAGG 1581
QY 1330 CATGGAGGAGCTCGCTCGGAGACCGTTCGACCATCAAGAAAGTGTGATCGCGGACA 1389
Db 1582 CATGGAGGAGCTCACTGGAGACCGTTCGACCATCAAAAGCTGTGATCGCGGACA 1641
QY 1390 GGATCAGGCGCTCGAGCTCGAGAAAGTGCCTGCGAGAAAGTAGGTGAGAGCGCCTCACA 1449
Db 1642 GGATCAGGCGCTCGAGCTCGAGAAAGTGCCTGCGAGAAAGTAGGTGAGAGCGCCTCACA 1701
QY 1450 AGAAGATGAGGAGACCATCGAGAGCTGAGACCGCATCGTTAAACAGCTCTCTCCATG 1509
Db 1702 AGAAGATGAGGAGCTCATCGAGAGCTGAGACCGCATAGTGAACAGCTCTCTCAG 1761
QY 1510 GCCCGCTGCACACCTGAGGTGCGACCGGACGACGACCGCCTTGACGAGAGCGCTCG 1569
Db 1762 GCCCATTCGACACCTGAGATGCGACCGGACGACGCGCACCTTGGACGAAACACTCG 1821
QY 1570 AGAATGACGACCGCTCAACCGGATGTTTACGCTCGACATCGAGAGCGCATCTCGAGC 1629
Db 1822 AGAATGACGACCGCTCAACAGAAATGTTTACGCTCGACATCGAGAGCGCTCTCTCGAGC 1881
QY 1630 AGAAGATCAAGCCCAAGGTGGAGAAAGACACAAAACCTGAGGCGCAGGAGCAATTTTC 1686
Db 1882 AGAAGATCAAGCCCAAGGTGGAGAAAGACACAAAACCTGAGACGAGAGCAGTTGTTC 1938

RESULT 5

D88382
LOCUS D88382
DEFINITION Hordeum vulgare hemA1 mRNA for glutamyl-tRNA reductase, complete cds.
ACCESSION D88382
VERSION D88382.1 GI:2967440
KEYWORDS hemA1; glutamyl-tRNA reductase.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Hordeum.
REFERENCE 1 (sites)
AUTHORS Tanaka, R., Yoshida, K., Nakayashiki, T., Tsuji, H., Inokuchi, H., Okada, K. and Tanaka, A.
TITLE The third member of the hemA gene family encoding glutamyl-tRNA reductase is primarily expressed in roots in Hordeum vulgare
JOURNAL Photosyn. Res. 53, 161-171 (1997)
AUTHORS Tanaka, R.
TITLE Direct Submission
JOURNAL Submitted (13-OCT-1996) Ryouichi Tanaka, Kyoto University, Department of Botany, Faculty of Science, Sakyo-Ku, Kitashirakawa-oiwakecho, Kyoto, Kyoto 606-01, Japan (E-mail: ryou@ok-lab.bot.kyoto-u.ac.jp, Tel: 075-753-4147, Fax: 075-753-4122)
FEATURES
source Location/Qualifiers
1..1864
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/sub_species="bonus"
/db_xref="taxon:112509"
/clone="BHA100"
1..1864
/gene="hemA1"
16..1599
/gene="hemA1"
/note="isoform 1"
/codon_start=1
/product="glutamyl-tRNA reductase"
/protein_id="BAA25167.1"
/db_xref="GI:2967441"
/translation="MAGATATAAGAAFAAKAGPAAACPMVLVAAGRRRSQVVRCD AGDAQNAASKAASITALEQFKISADRYMKESSTAVIGLSVHTAPVEMREKLVAEEL

REFERENCE 1
AUTHORS Cahoon,R.E., Gutteridge,S., Harvell,L.T., Rafalski,J.A., Tao,Y. and Weng,Z.
TITLE Polynucleotides encoding aminolevulinic acid biosynthetic enzymes
JOURNAL Patent: WO 0109304-A 7 08-FEB-2001;
E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES Location/Qualifiers
source 1..1778
/organism="Oryza sativa"
/mol_type="genomic DNA"
/db_xref="taxon:4530"
BASE COUNT 496 a 386 c 479 g 417 t
ORIGIN
Query Match 62.2%; Score 1196.4; DB 6; Length 1778;
Best Local Similarity 87.0%; Pred. No. 1.3e-195;
Matches 1314; Conservative 0; Mismatches 196; Indels 0; Gaps 0;
QY 210 GCCGCGTGGAGGCCAGGCCAGCCGCTGCGCAAGGCTGCCAGGTCGCCGCCCTCGAG 269
DB 1 GCACAGTGGTACCCAGCGCGAGCGGTGGCCAAAGGCCGCGCTGCCGCGCTCGAG 60
QY 270 CAGTTCAAGATATCCGCGACCGGTACATGAAGAAAGGAGTACCATAGCTGTGATAGCC 329
DB 61 CAGTTCAAGATCTCCGCGACCGGTACATGAAGAAAGGAGTACGATAGCGGTATAGCC 120
QY 330 CTCAGTGTACACAGCACAGTGGAGATGGGTGAAAACCTTGCTTTGCAGAGAACTG 389
DB 121 CTCAGTGTACACACTGCACCGAGTGGAGATGGGTGAGAACTTGCTTTGCAGAGAACTA 180
QY 390 TGGCCCGCTGCTATTCAAGAACTCACTAGCTGCAACCATATTGAAGGCTGCTGTTCTT 449
DB 181 TGGCCCGCTGCTATTCTAGAACTCACAGTCTGAATCATATTGAAGGCTGCTGTTCTT 240
QY 450 AGTACCTGAATAGAAATTTATGTGGCGCTTATCATGGAACCGTGGTATCAGA 509
DB 241 AGTACCTGAATAGAAATCTATGTGTAGCTTTATCGTGAACCGTGGGATTAGA 300
QY 510 GAAGTAGTGGACTGGATCTCGAAGAAAGTGGTATTCGGCTTCGGAGCTTAGGAGCAC 569
DB 301 GAAGTGGTAGACTGGATCTCAAGAAAGAGTGGAAATCCCTGCTCTGAGCTCAGGAGCAT 360
QY 570 CTGTTTCATCTTCGGAAGCAGTGATGCCACAGCCATCTGTTTGAAGTGTGAGCTGGCCTT 629
DB 361 CTATTCATGTCGTCGACAGTGATGCCACAGCCATCTGTTGAGGTAICTGCTGGGCTT 420
QY 630 GACTCTTTGTTCTCGGTGAAGGACAAATCTTGTCTCAGGTTAAACAAAGTTGTGAGGAGT 689
DB 421 GACTCTTTGTTCTTCGGAAGGAGGCAATCTTGTCTCAAGTTAAACAAAGTTGTGAGGAGT 480
QY 690 GGACAGACACTGGAGGCTTCGGAAGAACATCGATAGATGTTCAAGGATGCAATCACT 749
DB 481 GGGCAAAACAGTGGAGGCTTCGGAAGAACATCGATAGATGTTCAAGGATGCAATCACT 540
QY 750 GCTGGAAGCGTGTCCGCGAGCAGACCAACATATCATCTGCTGCTGTTTCTGTCAGTTCA 809
DB 541 GCTGGAAGCGTGTCCGCTGCGAGACTAACATATCATCAGTGTGCTCTCTGTCAAGTTCA 600
QY 810 CGCGGGTGTAACTGGCCCTGATGAAGCTTCGGAAGTCTGAAGCACTGTCAAGTAGGATG 869
DB 601 CTTGAGTGTAAATGGGCTTCGGAAGGCTTCGGAAGGCTTCGGAAGGCTTCGGAAGGCTTC 660
QY 870 CTTCTGATTTGCTGGTAAATGGGAGCTAGTGTCAACATCTGTTGTTCCCAAGGA 929
DB 661 CTGTTGATTTGCTGGCAGATGGGAAAGTTGGTGGTTAAACATTTGATTGCCAAGGA 720
QY 930 TGCAGAAGGTTGTTGTTGGTGAACCGCTTCGTTGGAAGGTTGGATGCTATTCTGTCAGGAG 989
DB 721 TGCAGAAGGTTGTTGTTGGTGAACCGCTTCAGTGGAAAGGTTGGATGCCATCCGCGAAGAG 780
QY 990 ATGAAGATATAGAGATCGTGTACAGGCTCTCTCAGACATGTATCAAGCTGCTGCTGAA 1049

Db 781 ATGAAGACATTTGAGATTGTGTACAGGCTCTTACAGAGATGTATGAAGCCGCTGCCGNA 840
QY 1050 GCTGATGTGCGTTTCAACAGCACCGCATCTGAAACTTCAATTTGTCGAAAAAGAACACGCA 1109
Db 841 GCTGATGTGCGTTTCAACAGCACCGCATCTGAAACTTCAATTTGTCGAAAAAGAACACGCA 900
QY 1110 GAGCAGCTCCCGCTGCTCTGATACCTATGAGAGGTTGTCGCTGTTGTCGACATATCT 1169
Db 901 GAGCGCTTCCCGCTATTTCTGATGCTATGGTGGTGTTCGACTCTTTGTGACATATCC 960
QY 1170 GTCCCCCAGGAATGTCAAGCATGTGTCTGAAGTTGGCGCTGCACGAGTGTACAAATGTC 1229
Db 961 GTCCCCCAGGAATGTCAAGCATGTGTCTGAAGTTGGCGCATGGCGAGTATACAACTGTC 1020
QY 1230 GACGACTTGAAGAGGTGGTGAAGCAACAAAGAGGAGCCGCTCAGAAAAAGCAATGAG 1289
Db 1021 GATGACTTGAAGAGGTGGTGAAGCAACAAAGAGGAGCCGCTTAGGAAAGCAATGAG 1080
QY 1290 GCGCAGACATCATCACCGAAGACTGAGAGCGTTTCGAGGATGAGAGGACTCGCTGAG 1349
Db 1081 GCCCAACAAATCATCACCGAAGACTGAAACCGTTTCGAGGATGAGAGGACTCGCTGAG 1140
QY 1350 ACCGTTCCGACCATCAAGAACTGAGGTCTGACGCGACAGATCAGGCGCTTCGAGGCTC 1409
Db 1141 ACTGTTCCGACTATCAAGAACTGAGGTCCTACGCGACAGATCAGGCGCTTCGAGGCTT 1200
QY 1410 GAGAAGTGCCTGCAAGAAAGTAGGTGAGGAGCCCTTCACCAAGAAAGATGAGGAGAGCCATC 1469
Db 1201 GAGAAGTGCCTTCCAGAAAGATCGGCGAAGAGCGCCCTCACCAAGAAAGATGAGGAGATCCATC 1260
QY 1470 GAGAGCTGAGCAGCCGATCGTTAAACAGCTCTCTCATGCGCGCTGCAGGACTGAGG 1529
Db 1261 GAGAGCTGAGCAGCCGATCGTTAAACAGCTCTCTCATGCGCGCTGCAGGACTGAGG 1320
QY 1530 TGCACCGCAGCAGCAGCCGATCGTTAAACAGCTCTCTCATGCGCGCTGCAGGACTGAGG 1589
Db 1321 TGTGACGCGCAGCAGCAGCCGATCGTTAAACAGCTCTCTCATGCGCGCTGCAGGACTGAGG 1380
QY 1590 CGGATGTTGAGCTTCGACATGAGAGAGCGGATCATTCGAGCAGAGATCAAGGCCAAGGTG 1649
Db 1381 AGGATGTTGAGCTTCGACATGAGAGAGCGGATCATTCGAGCAGAGATCAAGGCCAAGGTG 1440
QY 1650 GAGAGACACAAACTGAGGCGCAGGAGCAATTTTCTTACCACCATATCTATATATATA 1709
Db 1441 GAGAAGTCCCAAGAACTGAGATTGAAGAGAGATTTTTTTTTTTCAGCCCGTGTATCTACT 1500
QY 1710 GCGTCTCCAA 1719
Db 1501 ATGTATACTA 1510
RESULT 7
AX081089
LOCUS Sequence 19 from Patent WO0109304.
DEFINITION AX081089
ACCESSION AX081089
VERSION AX081089.1 GI:13170008
KEYWORDS
SOURCE
ORGANISM
Triticum aestivum (bread wheat)
Triticum aestivum
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
REFERENCE 1
AUTHORS Cahoon,R.E., Gutteridge,S., Harvell,L.T., Rafalski,J.A., Tao,Y. and Weng,Z.
TITLE Polynucleotides encoding aminolevulinic acid biosynthetic enzymes
JOURNAL Patent: WO 0109304-A 19 08-FEB-2001;
E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES Location/Qualifiers
source 1..1480
/organism="Triticum aestivum"
/mol_type="genomic DNA"

Db	966	AGGAAGACCGTGTGAGGAAGCAATGAGGCGCCAAACAATCATTTACCAAGAACTGAAC	1025
Qy	1321	GGTTCGAGGCGATCGAGGGACTCGCTCGAGACCGTTTCGACCATCAAGAGCTGAGTCTG	1380
Db	1026	GGTTCGAGGCGATCGAGGGACTCACTCGAGACCGTTTCGACCATCAAAAAGCTGAGTCTG	1085
Qy	1381	ACGGGACAGGATCAGGGCGCTCCGAGCTCGAGAAAGTGCCTCGAGAAAGTAGGTGAGGACG	1440
Db	1086	ACGGGACAGGATCAGGGCGATCCGAGCTCGAGAAAGTGTCTGCAAGAGATCGGGGAAGACA	1145
Qy	1441	CCCTCAACAAAGATGAGGAGAGCCATCGAGGAGCTGAGCACCGGCATCGTTTAAACAAGC	1500
Db	1146	ATCTCAACAAAGATGAGAGGTTCCATCGAGGAGCTGAGCACCGGCATAGTGAAACAAGC	1205
Qy	1501	TCCTCCATGCGCGCTCGAGGACCTGAGTGCAGCGGAGGACGACCGGACCCCTTGAGC	1560
Db	1206	TCCTTCACGCGCCCACTGCGAGCACCTGAGATGCGACGGGACGACCGCACCTTGAGC	1265
Qy	1561	AGACGCTCGAGAAACATGCGACCGCCTCAAACCGGATGTTTCAGCCCTCGACATGGAGAGCGCA	1620
Db	1266	AAACGCTTGAGACATGCGACCGCCTCAAAGATGTTTCAACCTCGACACGGAGAGCGCG	1325
Qy	1621	TCATCGAGCAGAAGATCAAGGCCAAGGTGAGGAGACACAAACCTGAGGCCAGGAAAGCAA	1680
Db	1326	TCCTTGAGCAGAAGATCAAGGCCAAGGTGAGGAGACACCAAGCTGAGACCGAGGACAC	1385
Qy	1681	TTTTTCTACACCAATTCATCTATATATAGGCTCTCAATCTCATTCATTTT	1733
Db	1386	TTGCGCGTC--TGTATATCTACTTATCTACTGCTGCCAGAAATGCGCTACATTTCT	1436
RESULT 8			
D88383			
LOCUS	D88383	Hordeum vulgare hema3 mRNA for glutamyl-L-cysteine synthetase, complete cds.	1822 bp mRNA linear PLN 17-MAR-1998
DEFINITION			
ACCESSION	D88383		
VERSION	D88383.1	GI:2967442	
KEYWORDS	hema3; glutamyl-L-cysteine synthetase.		
SOURCE	Hordeum vulgare subsp. vulgare		
ORGANISM	Hordeum vulgare subsp. vulgare		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Hordeum.		
REFERENCE	1 (bases 1 to 1822)		
AUTHORS	Tanaka, R., Yoshida, K., Nakayashiki, T., Tsuji, H., Inokuchi, H., Okada, K. and Tanaka, A.		
TITLE	The third member of the hema gene family encoding glutamyl-L-cysteine synthetase is primarily expressed in roots in Hordeum vulgare		
JOURNAL	Photosyn. Res. 53, 161-171 (1997)		
REFERENCE	2 (bases 1 to 1822)		
AUTHORS	Tanaka, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (13-OCT-1996) Ryouichi Tanaka, Kyoto University, Department of Botany, Faculty of Science, Sakyo-ku, Kitashirakawa-oiwakecho, Kyoto, Kyoto 606-01, Japan		
	(E-mail: ryouichi@lab.bot.kyoto-u.ac.jp. Tel: 075-753-4147, Fax: 075-753-4122)		
FEATURES			
source	Location/Qualifiers		
	1..1822		
	/organism="Hordeum vulgare subsp. vulgare"		
	/mol_type="mRNA"		
	/sub_species="bonus"		
	/db_xref="taxon:112509"		
	/clone="BHA200"		
gene	1..1822		
	/gene="hema3"		
CDS	28..1635		
	/gene="hema3"		
	/note="isoform 3"		
	/codon_start=1		
	/product="glutamyl-L-cysteine synthetase"		

TITLE
 Members of a low-copy number gene family encoding glutamyl-tRNA
 reductase are differentially expressed in barley
 Plant J. 9 (6), 867-878 (1996)
REFERENCE
 MEDLINE
 96267576
 8696365
REFERENCE
 2 (bases 1 to 1703)
 Bougri, O. and Grimm, B.
TITLE
 Members of a low-copy number gene family encoding glutamyl-tRNA
 reductase are differentially expressed in barley
 Unpublished
JOURNAL
 3 (bases 1 to 1703)
 Bougri, O.
REFERENCE
 Direct Submission
 Submitted (07-APR-1995) O. Bougri, Institute of Plant Genetics &
 Crop, Plant Research, Corrensstr. 3, 06466 Gatersleben, FRG
FEATURES
 Location/Qualifiers
 1..1703
 /organism="Hordeum vulgare subsp. vulgare"
 /mol_type="mRNA"
 /cui="112509"
 /db_xref="taxon:112509"
 /chromosome="5"
 /clone_lib="lambda ZAPII"
 /dev_stage="6 days old etiolated seedlings"
 1..1703
 /gene="hema 2"
 <1..1399
 /gene="hema 2"
 /function="aminolevulinate synthesis (chlorophyll
 synthesis)"
 /note="2nd isoform"
 /codon_start=2
 /product="glutamyl tRNA reductase"
 /protein_id="CAA0055.1"
 /db_xref="GI:1666079"
 /db_xref="SWISS-PROT:Q96563"
 /translation="QFKISADRYIKKSSIAVIGLSVHTAPVDMREKLAVABELMPRA
 ISLSTLSNHLIEAAVLSTGNRMETVYVALSWNRIGREVVDWMSKSGIPASELREHLF
 MLRDSGATRIHFEVSAGLSLVLGEGOILAQKVQVWNGQSGGLKRNIDRMPKDAIT
 AGKRCATRIHFEVSAGLSVSSAAVELAMKLPKSECLARMLLIAGKRGKLVVKKHLTA
 KGCKVVVRSVERVDAREMKDIEIVRPUTEMYEAADADVVFSTASLSLFT
 KEHAELPPLISLAWGVRLFDVSPRNVGACLSQVHEARVYVNDLKEVEANKEDR
 VRKAMEAQAIIITOLKRFPEARWDSLETPTIKLRSYADRIRASLDELKQIKI GEONL
 NKTRRSIEBELSTGIVNKLHLPLOHLRCDSGDSRTLDELTDNHNALNRNFNLDEKA
 VLEQIKIKAVEKQS"
 BASE COUNT 466 a 393 c 463 g 381 t
 ORIGIN
 Query Match 56.7%; Score 1090.6; DB 8; Length 1703;
 Best Local Similarity 84.5%; Pred. No. 1.9e-177;
 Matches 1225; Conservative 0; Mismatches 224; Indels 0; Gaps 0;
 QY 269 GCAGTTCAAGATATCCGCCACCGTACATGAGGAAAGAGTACCATAGCTGTATAGG 328
 DB 1 GCAGTTCAAGATCTCCGCACACCGTACATCAAGGAAAGAGTACCATAGCTGTATAGG 60
 QY 329 CCTCAGTGTACACAGCACACGAGTGGAGATGCGTGAAGAACTTCTCTTCAGAGGAACT 398
 DB 61 CCTCAGTGTACACAGCGCGGTGGACATGCGTGAAGAACTTCTCTTCAGAGGAACT 120
 QY 389 GTGGCCCGGTGCTATTCAAGAACTCACTAGCCCTGAACCATATTGAAGAGGCTGCTGTCT 448
 DB 121 ATGGCCCGGTGCTATTTCAGAACTCACCACTGTAACCATCGAAGAGGCTGCTGTCT 180
 QY 449 TAGTACTGTAATAGATGAATATTTATGTGGTGGCGCTATCATGGAACCGTGTATCAG 508
 DB 181 TAGTACTGTAATAGATGAATATTTATGTGGTGGCGCTATCATGGAACCGTGTATCAG 240
 QY 509 AGAAGTAGTGCATCTGGATGTCGAAGAAAGTGTATTCCCGCTTCCGAGCTTAGGGAGCA 568
 DB 241 AGAAGTAGTGCATCTGGATGTCGAAGAAAGTGTATTCCCGCTTCCGAGCTTAGGGAGCA 300
 QY 569 CTTGTTCACTCTCGAAGCAGTGAATGCCACACGCCCATCTGTTGAGGTGTCACTGCGCT 628

Db 301 TCCTTTATGTTGGTGCACAGTGTGTGCACACGCCATCTGTTGAGGTATCGCTGGCT 360
 QY 629 TGACTCTTTGGTTCCTCGGTGAAGCAAAATCCTTGTCTCAGGTTAAACAAGTTGTGAGAG 688
 Db 361 CGACTCTTTGGTTCCTCGGTGAAGCAAAATCCTTGTCTCAGGTTAAACAAGTTGTGAGAA 420
 QY 689 TGGACAAACAGTGGAGGCTTGGGAAAGAACATCATAGGATGTTCAAGGATGCATCAC 748
 Db 421 TGGGCAAAATAGTGGAGGCTTGGGAAAGAACATCATAGGATGTTCAAGGATGCATCAC 480
 QY 749 TGCTGAAAGCGTGTCCGACGAGAGCAACATATCATCTGGTCTGTTCTTGTCTCAGTTC 808
 Db 481 GGCTGGAAGCGCGCCGCTCGGAGCAACATATCATAGTGTGTCGCTGCTGTCTGAGTTC 540
 QY 809 AGCGCGGTTGAATCGCCCTCGATGAAGCTTCGGAAGTCTGGAAGCACTGTGAGTGGAT 868
 Db 541 AGCGCGAGTTGAGTTGGCCATGATGAAGCTTCGAAGTCCGAATGCTTGTGAGGAGGAT 600
 QY 869 GCTTCTGATTTGGTGTCTGTAATAATGGGAAAGCTAGTATCAACATCTGTTGTCAGAAAG 928
 Db 601 GCTTTTGAATTTGGCGCCGCGCAAAATGGGGAATTTAGTGTCTCAACATTTGATTGCAAAAG 660
 QY 929 ATGCAAGAAAGTGTGTGTGTGTGAACCGCTCCGTGGAAGGCTGATGCTATTTCGTGAGCA 988
 Db 661 ATGCAAGAAAGTGTGTGTGTGTGAACCGCTCCGTGGAAGGCTGATGCTATTTCGTGAGCA 720
 QY 989 GATGAAAGATATAGAGATCGTGTACAGGCTCTCTCAGACATGTATCAAGCTGTGCTGTA 1048
 Db 721 GATGAAAGATATAGAGATCGTGTACAGGCTCTCTCAGACATGTATCAAGGCTGTGCTGTA 780
 QY 1049 AGCTGATGTCTGTTTACAGACCGCATCTGAAACTTCTATTGTTGCGAAAGAACACGC 1108
 Db 781 TGCCGACGTTGTGTTTACAGACCGCATCTGAATCCCTATTATTCACCAAGGAGCATGC 840
 QY 1109 AGAGGCACTCCCGCTGTCTGATATCTATGGGAGGTGTCGCTGTTTGTGCGACATATC 1168
 Db 841 AGAGGCGCTTCTCTTATCTCTTGGCATGGGTGGTGTGCGCTTTTGTGCGACATATC 900
 QY 1169 TGTCCCGAGGAATGTTCAGCGCATGTGTGTCTGAAGTTTGGCGCTGCACGAGTGTCAATGT 1228
 Db 901 CGTCCCAAGGAAAGTGTGCGTGCCTTCTCAAGTGGAGCACGACGCGGTATACAAATGT 960
 QY 1229 CGAGCACTTGAAGAGGTGTGGAAGCAACAGAGGAGACCGGTCTAGGAAAGCAATGGA 1288
 Db 961 TGACGACTTGAAGAGGTGTGGAAGCAACAGAGGAGACCGGTCTAGGAAAGCAATGGA 1020
 QY 1289 GCGCGACAGCAATCATCAACGAGAACTGACACGCTTCGAGGCACTGGAGGAGCTCGCTGGA 1348
 Db 1021 GGCCCAAGCAATCATCAACGAGAACTGGAAGCGTTCGAGGCGTGGAGGAGCTCGCTGGA 1080
 QY 1349 GACCGTTCCGACCATCAAGAACTGAGGTGCTACGCGGACAGGATCAGGCGCTCGAGCT 1408
 Db 1081 GACCGTTCCGACCATCAAGAACTGAGGTGCTACGCGGACAGGATCAGAGCATCCGAGCT 1140
 QY 1409 CGAAGAGTGCCTCGAAGAGTAGTGAAGACGCTCTCAACAAAGATAGGAGAGCAT 1468
 Db 1141 CGAAGAGTGCCTCGAAGAGTAGTGAAGACGCTCTCAACAAAGATAGGAGAGCAT 1200
 QY 1469 CGAAGAGTGCCTCGAAGAGTAGTGAAGACGCTCTCAACAAAGATAGGAGAGCAT 1528
 Db 1201 CGAAGAGTGCCTCGAAGAGTAGTGAAGACGCTCTCAACAAAGATAGGAGAGCAT 1260
 QY 1529 GTGCGACGCGACGACGCGCACTTGAACGAGAGCTCGAGAACTGACGCGCTCAA 1588
 Db 1261 ATGCGATGCGACGCGACGCGCACTTGAACGAGAGCTCGAACAATGACGCGCTCAA 1320
 QY 1589 CCGGATGTTGAGCTCGAATGGAAGAGGAGTATCATGAGCAGAAATCAAGGCCAAGGT 1648
 Db 1321 CAGGATGTTCAACCTTTGACACGAGAGGCGGTCTCGAGCAGAAATCAAGGCCAAGGT 1380
 QY 1649 GGAAGAGACACAAACTGAGGCGCAGGAGCAATTTTCTACCACTTATCTATATATAT 1708

Db 1381 AGAGAAGCCAAAGCTGAGGACGAGGACGGTCCCTCGTCTGTATATCTACTTATAC 1440
 Qy 1709 AGCGTCTCC 1717
 Db 1441 CACCCACACC 1449

RESULT 10 AC078840/c

LOCUS AC078840 130058 bp DNA linear PLN 07-SEP-2002
 DEFINITION Oryza sativa chromosome 10 BAC OSJNB0073N24 genomic sequence,
 complete sequence.

ACCESSION AC078840
 VERSION AC078840.6 GI:22758306

KEYWORDS HTG.

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE AUTHORS

Buell, C.R., Yuan, Q., Moffat, K.S., Hill, J.N., Jenkins, C.N.,
 Burr, P.C., Hsiao, J., Zismann, V., Pai, G., Bowman, C.L., Fujii, C.Y.,
 VanAken, S.E., Bowman, C.L., Craven, B., Utterback, T.R., Khalak, H.,
 Feldblyum, T.V., Quackenbush, J., White, O., Salzberg, S.L. and
 Fraser, C.M.

TITLE Oryza sativa chromosome 10 BAC OSJNB0073N24 genomic sequence
 JOURNAL Unpublished

REFERENCE 2 (bases 1 to 130058)

AUTHORS Buell, R.

TITLE Direct Submission

JOURNAL Submitted (07-AUG-2000) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA

REFERENCE 3 (bases 1 to 130058)

AUTHORS Buell, R.

TITLE Direct Submission

JOURNAL Submitted (15-SEP-2000) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA, rbueller@tigr.org

REFERENCE 4 (bases 1 to 130058)

AUTHORS Buell, R.

TITLE Direct Submission

JOURNAL Submitted (22-SEP-2000) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA, rbueller@tigr.org

REFERENCE 5 (bases 1 to 130058)

AUTHORS Buell, R.

TITLE Direct Submission

JOURNAL Submitted (14-OCT-2000) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA, rbueller@tigr.org

REFERENCE 6 (bases 1 to 130058)

AUTHORS Buell, R.

TITLE Direct Submission

JOURNAL Submitted (07-SEP-2002) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA, rbueller@tigr.org

On Sep 7, 2002 this sequence version replaced gi:10140779.

COMMENT

Address all correspondence to:rice@tigr.org

BAC clone OSJNB0073N24 is from Oryza sativa chromosome 10
 The orientation of the sequence is from SP6 to T7 end of the BAC
 clone.

Genes were identified by a combination of several methods: Gene
 prediction programs including Genscan and Genscan+ (Chris Burge,
<http://CCR-081.mit.edu/GENSCAN.html>), GeneMarkHMM (Mark Borodovsky,
<http://genemark.biology.gatech.edu/GeneMark/>), and GeneSplicer
 (Michael Pertea and Steven Salzberg, contact.mpeptide.tigr.org),
 searches of the complete sequence against a peptide database and
 the plant EST database at TIGR (<http://www.tigr.org/tdb/tgi.shtml>).
 Annotated genes are named to indicate the level of evidence for
 their annotation. Genes with similarity to other proteins are named
 after the database hits. Genes without significant peptide
 similarity but with EST similarity are named as unknown proteins.
 Genes without protein or EST similarity, that are predicted by more
 than two gene prediction programs over most of their length are
 annotated as hypothetical proteins. Genes encoding tRNAs are

predicted by tRNAscan-SE (Sean Eddy,
<http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are
 identified by RepeatMasker (Arian Smit,
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>).

This BAC overlaps with rice BAC OSJNBa0078001 (AC079888) and
 OSJNBa0062C05 (AC084023).

FEATURES source

Location/Qualifiers

1. 130058

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="genomic DNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/chromosome="10"

/map="S21174S"

/clone="OSJNB0073N24"

/note="japonica cultivar-group"

/complement(311..1125)

/gene="OSJNB0073N24.23"

/note="predicted by genscan"

/complement(join(<311..452,944..>1125))

/gene="OSJNB0073N24.23"

/complement(join(311..452,944..1125))

/gene="OSJNB0073N24.23"

/codon_start=1

/product="hypothetical protein"

/protein_id="AAG13622.1"

/db_xref="GI:10140792"

/translations="MVMVTAALLLTDMAGVSYGRRLIPDLDAVVVGSGPPAKGGVM
 SRLQVPPSDSGHHVGDQNPYSQNFKVIDPKVFMAAQTLELQSLTRTTRISCKNSLFS
 QVGGK"

/complement(3145..3192)

/rpt_family="AT_rich"

/complement(3862..6664)

/gene="OSJNB0073N24.1"

/note="nearly identical to glutamyl tRNA reductase

GI:2920320 (Oryza sativa) (Plant Physiol 117, 332

(1998)); EST AU056848, D46763, AU056849, AU056803 from this

gene"

/complement(join(<3862..5094,5230..>5471,6430..>6664))

/gene="OSJNB0073N24.1"

/complement(join(3961..5094,5230..5471,6430..6664))

/gene="OSJNB0073N24.1"

/codon_start=1

/product="putative glutamyl-tRNA reductase"

/protein_id="AAG13620.1"

/db_xref="GI:10140790"

/translations="MASTTSATAGGAFAAKTRAGSSAAGGACARVAAGRRRSVG
 VPCDAGVEAQAOQAVAKAAVALEQFKISADRYMKERSIAVIGLSVHTAPVEMR
 EKLAVALBELPRAISELTSUNHIEAAVLSTCNMEIYVVALSWNRGIREVDMMSKK
 SGIPASBELRHLFMDRLSDATRHLEFVSAGLSVLGEGQILAQVKQVVRSGQSGGL
 GKNIDRMFKDAITAGKRVRCETNISSGAVSSAAVELALMKLPKSECLARMILLIGA
 GKMGKLVVXHLIAKGCKVNVNRSVERVDAREEMKDIIVYRPLTMYEAAAEADV
 VFSTASETPLTKHAELAPLSDAMGGVRLFDISVPNSVACVSEVGHARVYVND
 LKENVUANKEDRLKAMEAQTIITQELKRPFAWRSLSTVPTIKLRSADRIARASE
 LEKQKIGEDALTQKRRSIEELSTGVNKLHLGHLQHLRCDGSDSRITDLETLENMH
 ALNRMFSLDTEKAIIEQTKAKVQSKQN"

/complement(6480..6654)

/rpt_family="(CGG)n"

/complement(7437..7586)

/rpt_family="Galgin_Ol2 MITE element from gb:U72728 Oryza

longistaminata receptor-like kinase protein (Xa21), family

member F, pseudogene sequence (233 to 384) 152 nt"

/complement(7443..7584)

/rpt_family="Galgin_Os3 element from gb:D32165 Rice gene

for aspartic protease (302 to 448) 147 nt"

/complement(8050..8133)

/rpt_family="AT_rich"

/complement(8285..8320)

/rpt_family="AT_rich"

8873..8995

/rpt_family="Galgin_Os3 element from gb:D32165 Rice gene

for aspartic protease (302 to 448) 147 nt"

repeat_region	8873. .8994 /rpt_family="Gaigin_O12_MITE element from gb:U72728 Oryza longistaminata receptor-like kinase protein (xaz1), family member F, pseudogene sequence (233 to 384) 152 nt" complement(11172. .13291) /gene="OSJNBb0073N24.2" /note="similar to steroid membrane binding protein GI:1657409 (Sus scrofa); EST C26842,D48715, AA231657 from this gene" complement(join(<11172. .11511,12724. .13006,13085. .>13291)) /gene="OSJNBb0073N24.2" complement(join(11172. .11511,12724. .13006,13085. .13259)) /gene="OSJNBb0073N24.2" /codon_start=1 /product="putative steroid membrane binding protein" /protein_id="AAG13623.1" /db_xref="GI:10140793"	Db	5411	 TGGAGATCGTGAGAACTTGTGTGAGAGAACTATGGCCCCGTGCTATCTCAGAAC
gene		Qy	412	 TCACCTAGCCTGAACCATATTGAAGAGGCTGCTGTTCTTAGTACCTGTATAGAAAGGAAA
mrna		Db	5351	 TCACCAGTCTGAATCATATTAAGAGGCTGCTGTTCTTAGTACCTGCAATAGATGANA
CDS		Qy	472	 TTTATGTTGGTGGCGCTATCATGGAACCGTGGTATCAGAGAACTAGTGGAGTGTGCA
		Db	5291	 TCTATGTTGGTGGTATTTATCGTGGAAACCGTGGGATTAGAGAACTGGTGTAGACTGTGCA
		Qy	532	----- -----
		Db	5231	 AGGTAAGGATGTTTCATCAAAATGCTCATTTGTTCTGTCCATAGTCTTCTTTTCTCGTCT
		Qy	532	----- -----
		Db	5171	 TTTCCCTTTTCTGTGACTGAAGAACGGGTTTGTCTTAGATGGTTCTTAATCTCTGCT
		Qy	532	----- ----- AGAAAAGTGGTATTCCCGCTTCCGAGCTTAGGAGACCTGTTTCA
		Db	5111	 GCTTTTCTTCAACAGAAAAAGTGGAAATCCCTGCTTCTGAGCTCAGGGAGCATCTATTCA
repeat_region		Qy	577	 TCTTCGGAAGCAGTGCACACGCCATCTGTTTGGAGTGTGCTGAGCTGAGCTGCTGCTT
repeat_region		Db	5051	 TGTTCGCTGACAGTGTGCTGACGCCATCTGTTTGGAGTATCTGCTGGGCTTGACTCTT
gene		Qy	637	 TGGTTCTCGTGAAGGACAAATCTCTGCTCAGGTTTAAACAAGTGTGAGAGTGGACAGA
mrna		Db	4991	 TGGTTCTTGGAGAGGGCAATCTTCTGCTCAAGTTTAAACAAGTGTGAGAGTGGGCAAA
CDS		Qy	697	 ACAGTGGAGGCTTGGGAAGAACATCGATAGATGTTTCAAGATGCTCACTGCTGGAA
		Db	4931	 ACAGTGGAGGCTTGGGAAGAACATCGATAGATGTTTCAAGATGCTCACTGCTGGAA
		Qy	757	 AGCGTGTCCGAGGAGACCAATATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
		Db	4871	 AGCGTGTCCGCTGCGAGACTAACATATCATCAGGTGCTGCTGCTGCTGCTGCTGCTGCTG
		Qy	817	 TTGAACTGCGCTGATGAAGCTTCCGAAGTCTGGAAGCACTGTGCTGCTGCTGCTGCTGCTG
		Db	4811	 TTGAACTGCGCTGATGAAGCTTCCAAAGTCCGAATGCTTCTGCTGCTGCTGCTGCTGCTG
		Qy	877	 TTGCTGCTGTTAAATGGGAAGCTAGTGTGATCAAACTCTGCTGCTGCTGCTGCTGCTGCTG
		Db	4751	 TTGCTGCTGCAAGATGGGAAGTGGTGGTGTAAACAATTTCTGCTGCTGCTGCTGCTGCTG
		Qy	937	 AGGTTGTTGCTGAACCGCTCCGTTGGAAGGGTGGATGCTTATTCGTGAGGAGATGAAG
		Db	4691	 AAGTTGTTGTTGTAACCGTTCAAGTGGGAAGGGTGGATGCTTATTCGTGAGGAGATGAAG
		Qy	997	 ATATAGAGATCGTGTACAGGCTCTCTCAGACATGTATCAAGCTGCTGCTGCTGCTGCTGCTG
		Db	4631	 ACATTGAGATTGTGTAGAGGCTCTTACAGAGATGTATGAGAGCCGCTGCTGCTGCTGCTG
		Qy	1057	 TCGTGTTTACAGCACCGCATCTGAACTTCACTTGTGCTGCTGCTGCTGCTGCTGCTGCTG
		Db	4571	 TCGTGTTTACAGCACCGCATCTGAACTTCACTTGTGCTGCTGCTGCTGCTGCTGCTGCTG
		Qy	1117	 TCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
		Db	4511	 TCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
		Qy	1177	 GGAATGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
		Db	4451	 GGAATGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
		Qy	1237	 TGAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
		Db	4391	 TGAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
		Qy	1297	 CAATCATCCGGAAGACTGAGACGCTTCCAGGCACTGGAGGCACTGCTGCTGCTGCTGCTG

Db	4331	CAATCATCACCCAGAAATTGAAACGGTTTCGAGGATCGAGGACTCGCTGGAGACTGTTTC	4272
QY	1357	CGACCATCAAGAAGCTGAGTCTGACCGCGACAGATCAGGCGCTCGAGCTCGAGAAGT	1416
Db	4271	CGACTATCAAGAAGCTGAGTCTTACCGCGACAGATCAGGCGCTTCGAGCTCGAGAAGT	4212
QY	1417	GCCTGCAGAAAGTAGTTCAGAGCGCCCTCACCAAGAAGATGAGGAGCCATCGAGGAGC	1476
Db	4211	GCCTCCAGAAGATCGCGAAGACGCCCTCACCAAGAAGATGAGAAGATCCATCGAGGAGC	4152
QY	1477	TGAGCACCAGCGCATCGTTAAACAGCTCTCCATGGCCCGCTGCAGCACCTCGAGTGCAGC	1536
Db	4151	TCAGCACCAGCGCATCGTTAAACAGCTCTCCATGGCCCGCTGCAGCACCTCGAGTGCAGC	4092
QY	1537	GCAGCAGACGCGCACCTTCAGAGAGCGCTCGAGAACATGACCGCCCTCAACCGGATGT	1596
Db	4091	GCAGCAGACGCGCACCTTCAGAGAGCGCTCGAGAACATGACCGCCCTCAACCGGATGT	4032
QY	1597	TCAGCCTCGACATGAGAGAGCGGATCATCGAGCAGAGATCAAGCCCAAGTGGAGRAGA	1656
Db	4031	TCAGCCTCGACATGAGAGAGCGGATCATCGAGCAGAGATCAAGCCCAAGTGGAGRAGA	3972
QY	1657	CACAAACTGAGCGGAGGAGGAAAGCAATTTTCTACCAACCATTTATATATATAGCGTCTC	1716
Db	3971	CCCAGAACTGAGATGGAAGAGAGATTTTTTTTTTTCAGCCCGTGTATCTACTATGTATA	3912
QY	1717	CAA 1719	
Db	3911	CTA 3909	
RESULT 11			
AE017109/c			
LOCUS	AE017109	300029 bp	DNA linear PLN 06-JUN-2003
DEFINITION	Oryza sativa (japonica cultivar-group) chromosome 10, section 63 of 77 of the complete sequence.		
ACCESSION	AE017109	AE016959	
VERSION	AE017109.1	GI:31432860	
KEYWORDS	Oryza sativa (japonica cultivar-group)		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.		
REFERENCE	1 (bases 1 to 300029)		
AUTHORS	The Rice Chromosome 10 Sequencing Consortium		
CONSTRM	In-depth view of structure, activity, and evolution of rice chromosome 10		
TITLE	Science 300, 1566-1569 (2003)		
JOURNAL	Buell, C.R., Wing, R.A., McCombie, W.R., Messing, J. and Yuan, Q.		
AUTHORS	Direct Submission		
TITLE	Submitted (05-MAY-2003) The Institute for Genomic Research, 9712		
JOURNAL	Medical Center Dr, Rockville, MD 20850, USA		
COMMENT	This is the pseudomolecule for rice chromosome 10, which was constructed by resolving discrepancies between overlapping BACs, trimming the overlap regions, and linking the unique sequences to form a contiguous sequence. Genes in individual BAC clone were identified by a combination of several methods: Gene prediction programs, searches of the complete sequence against a peptide database and EST databases. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as unknown proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as hypothetical proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/).		
FEATURES	Location/Qualifiers		
source	1. 300029		
	/organism="Oryza sativa (japonica cultivar-group)"		
	/mol_type="genomic DNA"		

		misc_feature	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
		mRNA	
		CDS	
		gene	
	</		

```
SLIMKHFPSCFVVDNNYNTVNDLSDHLSGSSONKAHOVKIKOLRDHDLPLVAIIG
LLVNLVEXDLSNLRVLSGARVPDLSQNPQSEETORDIALLCSVFLASQGSASGT
ISPDEESLMOGAREAMWVEAVAAALLAELSTESMKVRGNAISSCLFNNSUKILVPA
LEKVSFHLQANMITEETHSAVEVIEKCKLS"
<19140. .>21483
/locus_tag="OSJNBa0017E08.6"
/note="similar to lysophosphatidic acid acyltransferase
GB:AAF73736 GI:8163563 (Brassica napus); EST C28836 from
this gene"
join(19140. .19391,19498.19683,19968. .20003,20294. .20700,
21048. .21149,21256. .>21483)
/locus_tag="OSJNBa0017E08.6"
join(19316. .19391,19498. .19683,19968. .20003,20294. .20700,
21048. .21149,21256. .21483)
/locus_tag="OSJNBa0017E08.6"
/codon_start=1
/product="putative lysophosphatidic acid acyltransferase"
/protein_id="AAP54440.1"
/db_xref="GI:31432864"
/translation="MGTLRPRPLAAGAGDATPSTAHAVVSGRGGRGVECOFHRV
RRRQPOVAVATASRRRTETVVRSDFAAGAAATMGDSQALSVSTPTWKVSMXYID
VVSVRGVCFYAVTAAIFLFAVMVHPLVLLDRYRRAQHYAKIWIATLISMIF
YKLDVEGHEMNPNSPVAIVANHOSFLDIYTLTLGRCKFKPISKTSIFMFPFIIGWM
YLLGPIRLRRMDSQLDCLKRCVLDVKKGASVFFPEGTRSKDGLGAFRGAFSA
TKTGAPVITPITLGTGKLMPSGMEGILNSGVKLIHHPIEGNDAEKLCSARKVIAD
TLILNGYGVH"
<25245. .>28542
/locus_tag="OSJNBa0017E08.21"
/note="similar to membrane associated protein Tb 291
GB:CAA52141 GI:393394 (Trypanosoma brucei)"
join(<25245. .25732,25802. .26164,26366. .26506,26602. .28303,
28396. .>28542)
/locus_tag="OSJNBa0017E08.21"
join(25245. .25732,25802. .26164,26366. .26506,26602. .28303,
28396. .28542)
/locus_tag="OSJNBa0017E08.21"
/codon_start=1
/product="putative membrane-associated protein"
/protein_id="AAP54441.1"
/db_xref="GI:31432865"
/translation="MARGSALLDGSVLPPSRIVSERQAGLPRRFPESATGREIVTLG
EGRPADPFGSRLFFPAMAGLVPFSSFFMDVLKFDYDLQMAHLTNAVNTLAFIAH
LCEMPVGRSLRFRWFFVTQSVSPVGGCYFQPMVPLNRYIPCALRKWDWK
SDCWRAPIQGLVDLRLAGLSQGLTGAMVGYDLRRRIAPLORARCAWEYITM
SEDTYRTHQVGRWQWADPDKFIVQVRLNLSMEASLIPOGILPCSDPDRASTLTG
TAVGASBERPGGAAPTQPEGERKKRLKMGETEPEGRNLISPPKSFNRPSEIIP
SRPSHSKSGSEADPAAAEARRRESDRRAARLREAEAEAEAEAEAEAEAEAEAE
EAAKARQAEAREADRAHQAERAEAREKFRQDEAMATSEARDEVAGASLEPAS
GDAQATSGAAGDEAGASLQTPGDAQDQGLRDIPESGTSGGSRVASSPRRLF
PTPSIAPLSAPFLQAALAAANIAVDGLSAQVEALQAEWAEALDAARVERGRSVEA
MVEVRKAHRRHVSELEARKKVLAEIAKEVEEREAGALIAIVSNVNEAODTLROYGSW
EAEGLKLDAAQGVLDVAAREQARGETEARSRRRETLERARNALEACVVERDLA
DREAUVTIREATLAHESACAEESALRHLHEDALTERALEAEAEAAQRLADLSLR
EAEQEQRATLECVRAERTALNQRAADLEAREKELDARSGGAAGLWENDLAARLAA
AEHTIADQQTLSNAGSEVEALRAGEVPGMLWDVAVSLDRAGRQGLWRGRTVKYA
ANHGLAQRLSKMAQLORLPEELEKTIKSSRLRDLQAQVELVRDAADHIVHSFEGB
PRLAFAPNSDEEDNAGADDSDDEAGDPGVD"
complement(<29170. .>35940)
/locus_tag="OSJNBa0017E08.20"
/note="similar to gag-poi precursor from gypsy type
retrotransposon GB:BAH84458 GI:5902445 (Oryza sativa)"
complement(join(<29170. .32047,32525. .33441,35108. .35540,
35663. .>35940))
/locus_tag="OSJNBa0017E08.20"
complement(join(29170. .32047,32525. .33441,35108. .35540,
35663. .35940))
/locus_tag="OSJNBa0017E08.20"
/codon_start=1
/product="putative gypsy-type retrotransposon"
/protein_id="AAP54442.1"
/db_xref="GI:31432866"
/translation="MRMGQTKFKSLRSEVTPWRCGBAADPAPPRPDPAIPRLDLAWR
RHEAARDGGGGSARCAVAARQARAGEVRNGMGLPRARRRSVAACGRVRRASCMPARPG
```

```
IGSRAAVVPLAASWAABRSOEGRDOAKGVQVGGAGAGATAYPALAOAMVVORELE
SMGCLPRCSRSPFLELIPFFCEBFLGNIETAAQRQKRLPKOCHVPFGSLAKTG
KAAGMNGVGLQSLGVLFLPSLMAQAKAHDLSVPSVGGDGPENRRRRTPTPLP
PQSPKGEALREVERSATPIAGDEERDRRLVYDGSSTPGQALQAAGALLRH
PPVVPDPESQARWLDDVAKLVMTVQRQDLAGRSSATKAGAAATGASRRRARH
AAAVHSAATSPSTPQKOLRGGDPARTSITERNRDSAAHATGASSSRVSPQHR
GNQPSVPVGGVGCRAFAVLNRNVRRPFRPTIAEKYDGSVNPBAEMLANPVVPKAN
GKLRMCIDYTLNACPKDPYLPRIIDQIVDSTAGRFISLGEKALPLKLRSGPP
TWYKXENALQLKAYLSSPPVLVAPDEPELLYLAATPQVVSAAALVVEREDNSHF
ALPHPPVPREQOGEAPEPNPGRPTTGVPLPACQTVLGAPOEGPEATVGRPK
HLSPPDEANPLTRPGKEQEEAPENGLRPLTTGVGLPACPTTGTGAPDQGPK
ATVGRPILLSDDPVISETDCAPRGCLDEECPRDAAPSEDRPHRKVRPVYVSEA
LRDAKRVPOAKMLYAILMASRKLRYHVFAHRTVTVTSYPLQGLNHRDNGTRVVKW
AIELSEFDLHFEPRHAIKSQALADFAVEMTAPPEVPSIPEASTDPSQLPHTAHWQF

Query Match      51.3%; Score 987; DB 8; Length 300029;
Best Local Similarity 79.5%; Pred. NO. 1.1e-159;
Matches 1243; Conservative 0; Mismatches 185; Indels 135; Gaps 1;

QY 292 GGTACATCAAGGAAGGAGTACCATAGCTGTGATAGGCTCAGTGTACACAGACACG 351
DB 264381 GGTACATCAAGGAAGGAGTACCATAGGCTCAGTGTACACAGACACG 264322
QY 352 TGGAGATCGGTGAAAAAAGCTCTGTGTGAGAGGAACTGTGGCCCGTCTATTCAAGAAC 411
DB 264321 TGGAGATCGGTGAGAAAGCTCTGTGTGAGAGGAACTATGGCCCGTCTATTCAAGAAC 264262
QY 412 TCATAGCTCAAGCATATTGAAGAGGCTGCTGTTCTTAGTACCTGTATAGAAAGGAAA 471
DB 264261 TCACAGCTCAATCATATTCAAGAGGCTGCTGTTCTTAGTACCTGCAATAGAAAGGAAA 264202
QY 472 TTTATGTGTCGGCGCTATCATCGGAACCGTGTGTATCAGAGAAGTAGTGACTGGATGTGCA 531
DB 264201 TCTATGTGTAGTCTTTATCGTGGAAACCGTGGGATTTAGAGAAAGTGTGTAGTGTGCA 264142
QY 532 ----- 531
DB 264141 AGGTAAGGATGTTTTCATCAAAATGCTCATTTGTTCTGTCCATAGTCTTCTTTCTCGTCT 264082
QY 532 ----- 531
DB 264081 TTTCCCTTTTCTGTACTGAAGAAACGGGTTTGTCTTAGATGGTTCTAATTCCTCTGCT 264022
QY 532 -----AGAAAAGTGGTATTCGGCTTCAGAGCTTAGGAGACCTGTTCA 576
DB 264021 GCTTTTTTTTCAACAGAAAAGTGGAAATCCCTGCTCTTCAGCTCAGGGAGCATCTATTCA 263962
QY 577 TCTTGGCAAGCAGTGATGCCACAGCCATCTGTTTGGAGTGTGACGTGCGCTTGGACTCTT 636
DB 263961 TGTTCGTGACAGTGATGCCACAGCCATCTGTTTGGAGTATCTGCTGGGCTTGACTCTT 263902
QY 637 TGGTTCTCGGTGAAGGACAAATCTCTGCTCAGGTTTAAACAAAGTTGTGAGGAGTGCACAGA 696
DB 263901 TGGTTCTCGGAGAGGACAAATCTCTGCTCAAGTTTAAACAAAGTTGTGAGAGTGGGCAA 263842
QY 697 ACAGTGGAGGCTGGGAAAGAACATCATCAGTAGGATGTTCAAGAGTGCATCATCTGCTGAAA 756
DB 263841 ACAGTGGAGGCTGGGAAAGAACATCATCAGTAGGATGTTCAAGAGTGCATCATCTGCTGAAA 263782
QY 757 AGCGTGTCCGAGGAGACCAACATATCATCTGCTGCTGCTGTTCTTCTGTCAGTTTCAGGGCGG 816
DB 263781 AGCGTGTCCGCTGGAGACTAAACATATCATCAGGTGCTGCTCTCTGTCAGTTTCAGCTGAG 263722
QY 817 TTGAATCGGCTTGGAGAGCTTCCGAAAGTCTGAAAGCACTGTGACGTAGGATGCTTCTGA 876
DB 263721 TTGAATCGGCTTGGAGAGCTTCCAAAGTCGGAATGCCATATCTGCTAGGATGCTGTTGA 263662
QY 877 TTGCTGCTGTAATAAGGAAAGCTAGTGATCAAAACATCTGTTTGCAGAAAGGATGCAAGA 936
DB 263661 TTGCTGCTGCAAGATGGGAAAGTTGGTGGTTTAAACATTTGATGCAAGGATGCAAGA 263602
QY 937 AGGTTGTTGGTGAACCGCTCCGTGGAAAGGTTGATGCTATTCTGTCGAGAGATGAAAG 996
```

```
Db 263601 AAGTTGTTGGTGAACCGTTTCAGTGGAAAGGTTGATGCCATCCGCGAAGAGATGAAG 263542
Qy 997 ATATAGAGATCGTGTACAGGCTCTCTCAGACATGTATCAAGCTGTGCTGAAGCTGATG 1056
Db 263541 ACATTGAGATTGTGTACAGGCTCTCTACAGAGATGTATGAAGCGCTGCGCAAGCTGATG 263482
Qy 1057 TCGGTTCACACGACGCGCATCTGAACCTTCATTGTTCCGAAAGACACACGACGAC 1116
Db 263481 TCGGTTCACAAGCAGCGCATCCGAAACCCCATTTGTTACAAAGAGACCGACGAGGCGC 263422
Qy 1117 TCCCGCTGTCTCTGATATGAGAGGTGTTTCGCTGTTTGTGACATATCTGTCCCA 1176
Db 263421 TTCCCGCTATTCTGATGCTATGGTGTGTTTCGACTCTTTGTGACATATCCGTCCCA 263362
Qy 1177 GGAATGTCAGCGATGTGTCTGAAGTTGGCGTGCACGATGTACAAATGTCAGAGCT 1236
Db 263361 GAAATGTCAGCGCTGTGTCTGAAGTTGGCGATGCGGAGTATACAACTGATGACT 263302
Qy 1237 TGAAGAGGTGTTGGAAGCCCAAGAGGAGGACCGCTCAGAAAGCAATGAGGCGCAGA 1296
Db 263301 TGAAGAGGTGTTGGAAGCCCAAGAGGAGGACCGCTTAGAAAGCAATGAGGCGCCAA 263242
Qy 1297 CAATCATCCAGAAAGCTGAGACGTTTCGAGGATGAGGAGCTCGCTGGAGACCGTTC 1356
Db 263241 CAATCATCCCAAGATTGAACGTTTCGAGGATGAGGAGCTCGCTGGAGACTGTTTC 263182
Qy 1357 CGACCATCAAGAGCTGAGTCTGTACGCGGACAGATCAGGCGCTCGAGCTCGAGAGT 1416
Db 263181 CGACTATCAAGAGCTGAGTCTGTACGCGGACAGATCAGGCGCTCGAGCTCGAGAGT 263122
Qy 1417 GCCTGCAGAAAGTAGTGAGGACCGCTCACCAAGAGATCAGAGAGCCATCGAGGAGC 1476
Db 263121 GCCTCCAGAGATCGCGAAGAGCGCTCACCAAGAGATCAGAGATCCATCGAGGAGC 263062
Qy 1477 TGAGCAGCGGATCGTTAAAGCTCTCCATGCGCCGCTGCAGCACCTGAGTGCGACG 1536
Db 263061 TCAGCAGCGGATCGTTAAAGCTCTCCATGCGCCGCTGCAGCACCTGAGTGAGC 263002
Qy 1537 GCAGCGACGCGGACCTTCAGCAGAGCTCGAGAGATCAGAGATCAGCGCTCAACCGATG 1596
Db 263001 GCAGCGACGCGGACCTTCAGCAGAGCTCGAGAGATCAGAGATCAGCGCTCAACCGATG 262942
Qy 1597 TCAGCCTCGACATGAGAGCGGATCATCGAGAGAGATCAAGCGCCAGGTGAGAGA 1656
Db 262941 TCAGCCTCGACATGAGAGCGGATCATCGAGAGAGATCAAGCGCGAGGTGAGAGA 262882
Qy 1657 CACAAACTGAGCGGAGAGAGCAATTTTCTACCAACCATATCTATATATAGCGTCTC 1716
Db 262881 CCCAGAACTGAGATTGAAGAGAGATTTTTTTTTTTCAGCCCGTGTACTACTATGATA 262822
Qy 1717 CAA 1719
Db 262821 CTA 262819

RESULT 12
AX081081
LOCUS AX081081
DEFINITION Sequence 11 from Patent WO0109304.
ACCESSION AX081081
VERSION AX081081.1 GI:13170004
KEYWORDS Glycine max (soybean)
SOURCE Glycine max
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1
REFERENCE
1
AUTHORS Cahoon,R.E., Gutteridge,S., Harvell,L.T., Rafalski,J.A., Tao,Y. and
Weng,Z.
TITLE Polynucleotides encoding aminolevulinic acid biosynthetic enzymes
JOURNAL Patent: WO 0109304-A 11 08-FEB-2001;
```

```
FEATURES
source
E.I. DU PONT DE NEMOURS AND COMPANY (US)
Location/Qualifiers
1..2055
/organism="Glycine max"
/mol_type="genomic DNA"
/db_xref="taxon:3847"
BASE COUNT 596 a 419 c 465 g 575 t
ORIGIN
Query Match 36.9%; Score 710.4; DB 6; Length 2055;
Best Local Similarity 68.1%; Pred. No. 4e-112;
Matches 1019; Conservative 0; Mismatches 471; Indels 6; Gaps 2;
Qy 189 GGGGTGTGGGTGCGGCTCGAGCGCGCGGGTGGAGGCGCCAGGCGCAGGCGCTGGCAAGGCT 248
Db 252 GGGGTATTTCGCTCGAGCGCTCGAGCGCTCTGATGCATCATCTGTGTCTCCAAATAATGCC 311
Qy 249 GCCAGCGCTCGCGCGCTCGAGCAGTTCGAAGATAT---CCGCGCAGCGGTACATGAAGAA 305
Db 312 ACCGCTCTCTCCGCTCTTGAGCAGCTCAAGNCTTCTGCAGCTGATAGATATACAAGAA 371
Qy 306 AGGATACCATAGCTGTGATAGGCTCAGTGTATACACAGCACCAGTGGAGATGCGTGA 365
Db 372 AGAAGCAGCATTATCGCATTTGGGCTCAGTGTGCACACTGCACCTGTGGAATGCGTGA 431
Qy 366 AAACCTGCTGTGAGAGGAACTGTGCGCGCGCTCTATTCAAGAACTCAGTACGCTGAAC 425
Db 432 AAACCTGCCATTCCAGAAGCAGAAATGGCTTAGAGCTATTGCGAGGCTGTGTAGTCTGA 491
Qy 426 CATATTCAAGAGGCTGTGTTCTTAGTACCTGTATAGATGGAATTTATGTGTGGCG 485
Db 492 CATATTGAAGAGCAGCTGTGCTGAGTACCTGCAATCGAATGGAGATATATGTTCTGCC 551
Qy 486 CTATCATGGAACCGTGTGTATCAGAGAAAGTAGTGGACTGGATGTCGAAGAAAGTGGTAT 545
Db 552 CTGTCCCAACATCGTGTGTCAAAGAGTCATGGAATGGATGTCAAAACAAGTCTCTGC 611
Qy 546 CCGCTTCCAGCTTAGGAGCAGCTGTTCATCTTGGAGCAGCTGATGCCACAGCCAT 605
Db 612 CCGTTCAGAGCTTAGCCAGCACCGGTTTTTACTTTACAAATGATGCCACACAGAT 671
Qy 606 CTGTTTCAGGTGTGAGTGGCTTGAATCTTTTGGTCTCGGTGAAGGACAAATCTTCTCT 665
Db 672 CTTTTCAGAGTATCAGCAGCTGTGACTCTCTGTTTGGGGGAAGGTCAAATCTTCTCT 731
Qy 666 CAGGTTAAACAAGTTGTGAGAGTGGACAGAACAGTGGAGGCTTTGGGAAGAATCGAT 725
Db 732 CAGGTTAAGCAAGTTGTTAAAGTTGGACAAGGAGTTAAACGCTTTGGGAGAAATATCAG 791
Qy 726 AGGATGTTCAAGGATGCAATCAGCTGGAAGAGGCTGTCGACGAGAGACCAATATCA 785
Db 792 GGGCTATTCAAGCATGCAATTTACTGTCGGGAAAGGGTTAGAACTGAGACTAATATTGCT 851
Qy 786 TCTGCTGCTGTTTCTGTCAGTTAGCGCGCGTTGAACTGGCCCTTGATGAAGCTTCCGAAG 845
Db 852 TCTGGGCGAGTTCTGTGAGCTCAGCTGCGGTGAGTTGGCTTATATGAGTTACTCTGAA 911
Qy 846 TCTGAAGCAGTGTGAGTGTCTCTGATTTGGTGTCTGTTGTTGTTGTTGTTGTTGTTG 905
Db 912 GCCTCACACGATAATGCGCAGATGTTGGTTATTGGTGTGTCGCAAGATGGGAAGCTTGTG 971
Qy 906 ATCAACATCTGTTGTCGAAGGATGCAAGAGGTTGTTGTTGTTGTTGTTGTTGTTGTTG 965
Db 972 ATCAACATTTGGTGGCAAGAGTTGCAAAAAGATGTTGTTGTTGTTGTTGTTGTTGTTG 1031
Qy 966 AGGTTGATGCTATTTCGTGAGGAGATGAAAGATATAGAGATCGTGTGACAGGCTCTCTCA 1025
Db 1032 AGAGTTGCTGCAATACGTGAAGAACTGAGAGATATGAGATATCTACAAACCCCTTCA 1091
Qy 1026 GACATGATCAAGCTGCTGCTGAAGCTGATGTGCTGTGTTCCACGACCGCATCTGAAACT 1085
Db 1092 GAAATGCTCACCTGTGTCGCGAAGCAGATTTAGTTTTTACCAGTACTGTCATCAGAAAC 1151
```


QY 1086 TCATTGTTGCAAAAGAACACGACGAGGACCTCCCTGCTCTGATACCTATGGAGGT 1145
 |||||
 Db 1152 CCATTATTTCTGAAAGAACATGTCAGGACCTTCTCTGCAAGTCAGAGATGGAGGC 1211
 |||||
 QY 1146 GTTCGCTGTTGTCGACATATCTGTCCCGAGGAATGTCAGGCAATGTGTCTGAAGTT 1205
 |||||
 Db 1212 CGTCGCTTTTTCATTGATATCTCTGTCCCGGAATGTGGGTTCATGTGTCTCAGACCTT 1271
 |||||
 QY 1206 GCGCTGACGAGTGTCAATGTCGACGACTTGAAGAGGTGTTGGAAGCCCAACAGAG 1265
 |||||
 Db 1272 GAGTCTGTGCGAGTTACATGTTGACGACCTTAAAGAGGTGTGGCTGCCAATAAGAG 1331
 |||||
 QY 1266 GACCGCTCAGAAAGCAATGAGCGCGCAGCAATCATCATCCGAGAACTGAGAGCGGTT 1325
 |||||
 Db 1332 GATCGCTTAAGAAAGCAATGGAAGCAGGCAATCATTTGCTGAAGATCTAAGCAATTC 1391
 |||||
 QY 1326 GAGCATGAGGAGTCTGCTCGAGACCGTTCGACCATCAAGAGCTGAGTGTGACGCG 1385
 |||||
 Db 1392 GAAGCTTGGAGGAGTCTACTCGAACTGTTCTACTATTAAGAAATGAGGGCTTATGCT 1451
 |||||
 QY 1386 GACAGGATCAGGCTCGAGCTCGAGAGTGCCTGCAAGAGTAGGTGAGGACGCTTC 1445
 |||||
 Db 1452 GAAGAAATCAGCTTGTGACCTTGAGAGTGTCTTAGTAAGATGGGTGATGA---TATA 1508
 |||||
 QY 1446 ACCAAGAGATGAGGAGGACCATCGAGGAGTGCAGCCGCTCGTTTAAAGCTCTC 1505
 |||||
 Db 1509 CCAAGAAACCGGAGAGCTGTGGATGACCTTAGTCGGGTATAGTGAATAGTGT 1568
 |||||
 QY 1506 CATGCCCGCTGACGACCTGAGTGGACGCGAGCAGCAGCCGACCTTGGACGAGCG 1565
 |||||
 Db 1569 CATGTCATGCAACATTTAAAGTGTGATGGAAACGACAGCCGAGCTCTTAGTGAGACA 1628
 |||||
 QY 1566 CTCGAGAAATGCAACGCTCAACCGGATGTTGAGCTTCGACATGAGAGGCGCATC 1625
 |||||
 Db 1629 CTGAGAACATGATGCTTTGATAGATGTTCAACCTTGAGACAGAAATATCTGTTTG 1688
 |||||
 QY 1626 GAGCAGAGATCAAGCCGAGGTCGAGAGACACAAACTGAGCCGAGGAGCAAT 1681
 |||||
 Db 1689 GAGGAGAAGATTTCGAGCAAGGTGCAACAAACAGAAATGAAATTAACACCAAT 1744
 |||||

RESULT 13

D50407
 LOCUS Cucumis sativus 2072 bp mRNA linear PLN 10-FEB-1999
 DEFINITION Cucumis mRNA for glutamyl-tRNA reductase, isozyme 1 complete cds.
 ACCESSION D50407
 VERSION D50407.1 GI:1694925
 KEYWORDS hema; glutamyl-tRNA reductase.
 SOURCE Cucumis sativus (cucumber)
 ORGANISM Cucumis sativus

REFERENCE
 AUTHORS Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.
 1 (bases 1 to 2072)

TITLE Tanaka, R., Yoshida, K., Nakayashiki, T., Masuda, T., Tsuji, H., Inokuchi, H. and Tanaka, A.
 Differential expression of two hema mRNAs encoding glutamyl-tRNA reductase proteins in greening cucumber seedlings

JOURNAL plant physiol. 110 (4), 1223-1230 (1996)
 MEDLINE 97088717
 PUBMED 8934625
 REFERENCE 2 (bases 1 to 2072)
 Tanaka, R.
 Direct Submission
 Submitted (23-APR-1995) Ryouichi Tanaka, Kyoto University, Department of Botany, Faculty of Science, Sakyo-ku, Kitashirakawa-oiwakecho, Kyoto, Kyoto 606-01, Japan (E-mail:ryouichi@ok-lab.bot.kyoto-u.ac.jp, Tel:075-753-4147, Fax:075-753-4257)

FEATURES
 Location/Qualifiers

1..2072
 /organism="Cucumis sativus"
 /mol_type="mRNA"

source

gene
 CDS
 /strain="cv. Aonagajibai"
 /db_xref="taxon:3659"
 /tissue_type="cotyledon"
 /dev_stage="greening"
 1..2072
 /gene="hema"
 78..1736
 /gene="hema"
 /codon_start=1
 /product="glutamyl-tRNA reductase"
 /protein_id="BAA08910.1"
 /db_xref="GI:1694926"
 /translation="MAVSTSFSGAKLRLFKSASNSSTNLSLSSHLPGFKSIRTR
 RILQRTGVSTFPFKCELASDVLVONDEIDPPKSNLSLEALQKTSADVRYTKERS
 SIVVIGLSIHITPEMEKLAIPAEPMFRAIGELCGLNHEEAAVLTSCNMEIYVVA
 LSQHRQVYKVTWMSKTSIPVBSICQHRFLVYNNDATQHI FEVSAGLDSLVLEGGQI
 LAQVQVYKVGQVAGFRNIGSLFKHATVTKVRVETETNIAAGVASVSSAAVEALAM
 KLPEPSHATARMLVIGAKMGKLVIKHLVAKGCTQWVNNRSBRTVAIREMKDVEI
 IYKPLTEMLSTAEADVIFTTSTASESLFTKEOVKDLPPVGHDDVGLRLFDLSVPRN
 VGACINLEEDRVYVNDLKEVVAANKEDRLRKAMEAQSIITEESKQFEAWRDSLET
 VTIKKRAYABRIARTAELEKLSRMGDDIPKTRRAVDDLSRGIVNKLHGPMLHLRC
 DGSDSRTLSLTLENMALRMFSLTETIAVLEQKIRAKVEQNQK"

BASE COUNT
 ORIGIN

Query Match 36.7%; Score 706.2; DB 8; Length 2072;
 Best Local Similarity 68.8%; Pred. No. 2.1e-111;
 Matches 1001; Conservative 0; Mismatches 448; Indels 6; Gaps 2;

QY 241 CAAAGGCTGCGAGCGTCCGCCCTCGAGCAGATTCAGAGATATCGC---CGACCGGTACA 297
 |||||
 Db 310 CTAATCTTCTATCTCTCGCCCTTGAACAGCTCAAGACCTCTGCTGTTCAGAGATATA 369
 |||||
 QY 298 TGAAGGAAGGAGTACCATAGCTGTGATAGCCCTCAGTGTACACAGCAGCAGTGGAGA 357
 |||||
 Db 370 CAAAGGAAGAGCAGCATTTGTTGTCATTGGGCTTAGTATTATACACACCTCTTGAAA 429
 |||||
 QY 358 TGGGTGAAAACCTGCTGTTGAGAGGAACTGGCCCCGCTGCTATTTCAGAACTCACTA 417
 |||||
 Db 430 TGAGAGAAAACCTGGCCATTCCTGAGCAGAGTGGCCCTCGAGCCATTGGAGAGCTCTGT 489
 |||||
 QY 418 GCCTGAACCATATTGAAGAGCGTCTGTTCTTAGTACTCTGTAATAGAAATGGAATTTATG 477
 |||||
 Db 490 GCTTAAACCATATTGAAGAGCAGCGCTTTTAGCACCCTGCAACAGAAATGAGATATATG 549
 |||||
 QY 478 TGTGGCGCTATCATGGAACCTGGTATCAGAGAAGTAGTGGAGTGTGCGAAGAAA 537
 |||||
 Db 550 TTGTTGCTTTGCTCAGCATCGTGGAGTGAAGGAGTGAAGTGAATGATGTCCAGACAA 609
 |||||
 QY 538 GTGGTATTCGCTTCGAGCTTAGGAGCACCTGTTTCATCTTTCGGAAGCAGTGTATGCCA 597
 |||||
 Db 610 GTGGAATCCCGGTTTCGGAGATTTCGAGCAGCACCGGTTTTTGTCTGTATAACAACGATGCCA 669
 |||||
 QY 598 CAGCCATCTGTTTGAAGGTGTGAGTGGCTTCAGCTCTCTTTGTTCTCGGTGGAAGCAAA 657
 |||||
 Db 670 CACAACATATTTTCGAAGTTTCAGAGGCTTCAGCTCTCTGTTCTTAGGAGAAGCCAAA 729
 |||||
 QY 658 TCCTTCTCAGGTTAAACAAAGTTGTGAGGAGTGGACAGAACAGTGGAGGCTTGGGAAAGA 717
 |||||
 Db 730 TCCTTCTCAGGTTAAACAAAGTTGTCAAGTTTCGACAGGCTGTTTCAGAGGATTTGGAAGGA 789
 |||||
 QY 718 ACATCATAGGATGTTCAAGGATGCAATCATCTCTGGAAGCGTGTCCGAGCGAGACCA 777
 |||||
 Db 790 ACATTAGTGGCTTTTTCAGAGCATGCCATCAGCTGGGTAAACCGGTGAAGAACAGAGACTA 849
 |||||
 QY 778 ACATATCATCTGCTGCTGTTTCTGTCAGTTTCAGCGCGGTTGAAGTGGCCCTCATGAGC 837
 |||||
 Db 850 ATATTCCTGCTGAGAGCTGTTTCTGAGTTCCTGCTCTGTCGAAATTTGGCCTTGTATGAGC 909
 |||||
 QY 838 TTCCGAAGTCTGAAGCAGTGTGAGTGTGCTGATAGGATGCTTCTGATTTGGTGTGCTGTAATAATGGGAA 897
 |||||
 Db 910 TTCTCTGAACCTTCACATGCCACTGCCAGAAATGTTGGTGTATGGAGCTGTGATAGATGGGAA 969
 |||||

Qy	898	AGCTAGTATCAAAACATCTGGTTGCTCCAAAGATGCAAGAAAGTTGTTGTGGTGAACCGCT	957
Db	970	AGCTTGTAAATTAACACATTTGGTTGCTAAAGCGCTGCACAAAAGATGGTTGTTGTGAATCGT	1029
Qy	958	CCGTGGAAAGCGGTGGATGCTATTTCGTGAGAGATGAAAGATATAGACATCGTGTACAGGC	1017
Db	1030	CTGNAGAGAGATCAAGACCATCGTGAAGAGATGAAGGATGTCGAGATTAATCTTACAGC	1089
Qy	1018	CTCTCTCAGACATGTATCAAGCTGCTCTGAAAGTGTATGTCGTGTTTACCAGCACCGCAT	1077
Db	1090	CCCTCACGGAAAATGCTTAGCTGTACAGCTGAAGCAGACGCTGATATTATACAGCACAGCTT	1149
Qy	1078	CTGAAACTTCATTTGTCGCAAAAGAAACACGACAGAGCACTCCCCCTGTCTCTGTACTA	1137
Db	1150	CTGAAAGTCTTTTGTATCAAAAGAGCAGGTTAAGGATCTTCCCTCTGTTGGCATGATG	1209
Qy	1138	TGGAGAGTGTTCGCTGTTGTGCACATATCTGTCCCCAGGAATGTCAGCGCATGTGTGT	1197
Db	1210	TTGGGGCTTGAGGCTTTTCATCGATATCTCTGTTCTCGAAGCTCGAGCATGTATCA	1269
Qy	1198	CTGAAGTTGGCGTGCACGAGTGTACAATGTGACGACTTTGAAAGAGGTTGGTGAAGCCA	1257
Db	1270	ACAACTCTGAAGATGTACGAGTGTACAATGTTGACGACCTTAAGGAGGTAGTTGCTGCAA	1329
Qy	1258	ACAAGGAGGACCGCTCTCAGAAAGCAATGAGGCGCAGACAATCATCACCGAAGAACTGA	1317
Db	1330	ACAAGGAAGATAGGCTCCGAAAGCAATGGAAGCAATCAATTTATCCTGGAAGAAATCCA	1389
Qy	1318	GACGGTTTCAGGCAATGAGGAGACTCGCTGGAGACCGTTTCGACCATCAAGAAAGCTGAGGT	1377
Db	1390	AGCAATTTGAAGCTTGGAGGATTTATGGAGACTGTTCCAAACAATTAAGAACTAAGGG	1449
Qy	1378	CGTACGGCGACAGATCAGGGCTCGGAGCTCGAAGAGTGCTGCGAAGAAAGTGTAGGTGAGG	1437
Db	1450	CTTATGCGCAAGAAATCAGAACTCGCGAGTTAGAAAATGCCTATCAAGAGATGGTGATG	1509
Qy	1438	ACGCCCTCACCAAGAATGAGGAGAGCCATCGAGGAGCTGAGCACCGGCATCGTTAAACA	1497
Db	1510	A---TATCCCAAGAAAGCTCGACAGCTGTGGATGATCTTAGTCTGGTATTAGTTAAACA	1566
Qy	1498	AGCTCTCCATTGGCCCTCGACACCTGAGGTGTCGACGGCAGCGACAGCGCACCCCTTG	1557
Db	1567	AGCTGCTTCATGTTGCTCTATGCAGCATTTGATGTCGACGGGAGTGACAGCCGAACCTTAA	1626
Qy	1558	ACGNAGAGCTCGAAGACATGACGCCCTCAACCGGATGTTACGCTTCGACATGGAGAAGG	1617
Db	1627	GTGAAACCTTTGGAGAAATCGACGCTCTAAATAGAATGTTTTCAGCTTGGAAACAGAATCG	1686
Qy	1618	CGATCATCGAGCAGAAGATCAAGGCCAAGGTGGAGAGACACAAAACTGAGGCGCAGGAAG	1677
Db	1687	CAGTGTGGAAACAGANAATCAGACTAAAGTGGNAACAAAACCAAGATGAAGCTAGTAGTAA	1746
Qy	1678	CAATTTTCTTACCAC	1692
Db	1747	TATTATATGAATCT	1761

RESULT	14
AX506506	
LOCUS	AX506506
DEFINITION	Sequence 1201 from Patent WO216655.
ACCESSION	AX506506
VERSION	AX506506.1 GI:23387743
KEYWORDS	.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi- s.
REFERENCE	1 Harper, J.F., Kreps, J.J., Wang, X. and Zhu, T. Stress-regulated genes of plants, transgenic plants containing same, and methods of use
AUTHORS	
TITLE	

JOURNAL Patent: WO 0216655-A 1201 28-FEB-2002; The Scripps Research Institute (US) ; Syngenta Participations AG (CH)									
FEATURES	source	Location/Qualifiers							
		1..1632 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /db_xref="taxon:3702"							
BASE COUNT	463 a	-289 c	428 g	452 t					
ORIGIN									
	Query Match	35.3%	Score 679.6;	DB 6;	Length 1632;				
	Best Local Similarity	68.8%	Pred. No. 7.9e-107;						
	Matches 985;	Conservative 0;	Mismatches 439;	Indels 12;	Gaps 3;				
Qy	235	CCGTGGCAAAAGGCTGCCAGGCTGCCGCCCTCCAGCAGATTCAAGATAT---	CCGCCGAC	291					
Db	206	CTGGCTAATGAGCTAGGATCTCTGCTTGACAACTCAGAAGCTCGCAGCTGATC	265						
Qy	292	GGTACATGAAGAAAGAGTACCATAGCTGTGATAGGCCTCAGTGACACAGCACCAG	351						
Db	266	GATATACAAAGAAAGAGCAGTATTGTTGTGATTGCACTTAGTATTCACACAGCTCCTG	325						
Qy	352	TGGAGATGCGTGAAGAACTTGCTGTTGACAGAGAACTGTGGCCCCCTGCTATTCAAGAAC	411						
Db	326	TTGAGATGCGGTGAGAAGCTTGCTATACCTGAAGCTGAATGGCCCAAGAGCTATTGCTGAAT	385						
Qy	412	TCACTAGCTGAACCATATTGAAGAGCGCTGCTCTTTAGTACTCTTAATAGAAATGGAAA	471						
Db	386	TGTGTGGTTGAATCATATTGAAGAAGCTGCTGTGCTTTAGTACTTGTAAACCGTATGGAGA	445						
Qy	472	TTTATGTGTGGCGCTATCATTGAACCGTGGTATCAGAAAGTAGTGGACTGATGTCGA	531						
Db	446	TTTTATGTTTGTAGCTTTATCTCAGCATCGTGGAGTTAAAGAAGTTACTGAATGCGATGCAA	505						
Qy	532	AGAAAGTGTGATTTCCGCTTCCGAGCTTAGGAGACACCTGTTTCATCTTCGCGAAGCAGTG	591						
Db	506	AGA CGAGTGGGATCCCAAGTTTCAGAGATTGTCAACACCGCTTTTCTGTTGTACAACAAGG	565						
Qy	592	ATGCCACAGCCATCTGTTTGGGTGTCAGCTGGCCCTTGACTCTCTTTGGTTCTCGGTGAAG	651						
Db	566	ATGCTACACAGCATATTTGAAGTCTCAGCTGGTCTTGACTCTCTTGCTTAGGAGAAG	625						
Qy	652	GACAAATCCTTGCTCAGGTATAA CAAGTTGTGAGGAGTGCAGAAACAGTGGAGGCTTGG	711						
Db	626	GT CAGATCCTTG CACAGGTGAAACAAGTTGTGAAAGTTGTTCAAGGAGTGAATGGCTTTC	685						
Qy	712	GAAAGAACATCGATAGATGTTTCAAGGATGCAATCACTGCTGGAAGAGCTGTCCGACGG	771						
Db	686	GGAGGAATATCAGCGGCTGTTTAAACACGCGATAACTGTTGGTAAAGCGTGTAGAACAG	745						
Qy	772	AGACCAACATATCATCTGGTGTCTTCTGTCAGTTTCAGCGGGTTGAACCTGGCCCTGA	831						
Db	746	AGACAATAATGCTTCTGGTGGGTTTCGGTTAGCTCAGCTGCTGTTTGAACCTTGCTCTGA	805						
Qy	832	TGAAGCTTCGGAAGTCTGAAGCACTGTCAGCTAGGATGCTTCTGATGGTGTGGTAAAA	891						
Db	806	TGAAGCTTCGCGAGTCTTCAACGATCTGCTAGGATGTGTAAATTTGGCGCTGGAAAAA	865						
Qy	892	TGGGAAAGCTAGTGATCAAA CATCTGGTTGCCAAAGGATGCAAGAGGTTGTTGTGGTGA	951						
Db	866	TGGGGAAGCTTGTGATTAAGCATTTTATGGCTTAAGGGTTGCACAAAAGTGGTGTAGTCA	925						
Qy	952	ACCGCTCCGTGGGAAAAGGGTGGATGCTATTCTGTTGAGGAGATGAAGATATAGAGATCGCT	1011						
Db	926	ACAAAGTGAAGAGAGGGTTTCAGCTATCCGTGAGGAGATGCTCGTATTTGAGATTTATAT	985						
Qy	1012	ACAGGCTCTCTCAGACATGTATCAAGCTGCTGCTGAAGCTGATGTCGTGTTTACCAGCA	1071						
Db	986	ATCGACCTCTTGATGATGATGCTAGCTTGTGCTTCAGAAAGGATGTTGTGTTTACAGCA	1045						
Qy	1072	CCGCATCTGAACCTTCATTTGTCGCAAAAGAACACGCCAGAGGCACTCCCCCTGTCCTG	1131						

Db 1046 CAGCCTCTGAGACGCCATTGTTCTTTAAAGGAGCATGTAGAGAAATCTCCCTCAAGCTTCTC 1105

Qy 1132 ATACTATGGAGGTGTTCGCTGTTTGTTCGACATATCTGTCCCGAGGAATGTCGCGCAT 1191

Db 1106 CAGAGTTGGAGGACTGAGGCAATTTTGTGACATCTCTGTTCCGGAATGTTGATCTT 1165

Qy 1192 GTGTCTCTGAAGTTGGCGCTGCACGAGTGTACAATGTTCGACGACTTGAAGAGGTGGTGG 1251

Db 1166 GTTGTGTGAAGTAGAACAACAGCAGCGGTTTACAATGTTCGACGATCTCAAGAGAGTTGTTG 1225

Qy 1252 AAGCCAAAGAGGACCGCTCAGGAAGCAATGAGCGCGGCGGACGACATCATCCCGAAG 1311

Db 1226 CTGCCAAACAAAGAACAGCAGGATGAGGAAGCAATGGAAGCTCAACCAATAATCACAGAGG 1285

Qy 1312 AACTGAGACGGTTTCAGGACATGAGGAGCTCCTCGAGACCGTTCCGACCATCAAGAGAC 1371

Db 1286 AATCCACACAGTTTGAAGCATGGAGGATTCATTGGAGACGGTTCCACGATCAAGAAGT 1345

Qy 1372 TGAGTCTGACCGACAGGATCAGGGCTCGAGCTCGAGAGTGCCTCGCAGAAAGTAG 1431

Db 1346 TAAGAGCTTATGACAGAGAAATCAGATGGCAGACTCGAGAAAGTGCATGTCCAAATGG 1405

Qy 1432 GTGAGGACCGCTCACCAGAGATGAGAGAGCATCGAGGAGCTGAGACCGGCATCG 1491

Db 1406 GAGATGA---TATCAACAAAGAAACAAACGAGAGCGGTTGATGACTTAAGCCGAGGTATCG 1462

Qy 1492 TTAACAAGCTCTCCATCGCCGCTGCGACACCTGAGGTGCGAGCGGACGACGCCCA 1551

Db 1463 TGAACAGATTCTTACATGCTCCATGCGACATTTGAGATGTGACGGAAGTGCAGTAGAA 1522

Qy 1552 CCCTTGACGAGCGCTCGAACAATGACGCGCTCAACCGGATGTTACGCTCGACATGG 1611

Db 1523 CGCTGAGCGAGACCTTTCAGAAATATGATGCTGTTTGAACAGAAATGTACGGTC-----TTG 1576

Qy 1612 AGAAGGCGATCATGAGCAGAGATCAGGCGCAAGTGGAGGAGCAGACCAAACTGA 1667

Db 1577 AGAAGACATTTTGAGGAAAGCTTAAGGCAATGGCGGCAACAAACAGAGATGA 1632

RESULT 15

LOCUS AY096600

DEFINITION Arabidopsis thaliana putative glutamyl-tRNA reductase (At1g58290)

ACCESSION AY096600

VERSION AY096600.1 GI:20465534

KEYWORDS FLI CDNA.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

AUTHORS Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

Arabidopsis Open Reading Frame (ORF) Clones

Unpublished

2 (bases 1 to 1663)

TITLE

JOURNAL Arabidopsis

AUTHORS Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

Direct Submission

TITLE

Arabidopsis

JOURNAL

COMMENT

Submitted (16-APR-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGSC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAPL cDNAs: Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGSC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGSC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

FEATURES

source

1. 1663

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="1"

/clone="U19478"

/note="This clone is in PUNI 51."

ecotype: Columbia"

gene

1. 1663

/gene="At1g58290"

CDS

1. 1632

/gene="At1g58290"

/codon_start=1

/evidence=experimental

/product="putative glutamyl-tRNA reductase"

/protein_id="AA020250.1"

/db_xref="GI:20465535"

/translation="MAVSAPVCPKLETLNHNHNSPSSSSSVSTPLGLNGVRVLPKNRTFRGLQKARCELSASDASNAASISALEOLKNSAARDYKERSIVVIGLSIHTAPVEMREKLAIPAEWPRAIAELGLNHIEAAVLSTCNREIIVLALSQRHGVKEVEMSKTSGIPVSEICORHFLYINKDATQHIPEVSAGLSDVLGEGQLAQKQVVKVQGVNGFRNIGLFIHATITVKEKRTETNIVASGVSAAVELALMKLPQSSVSARMCVIGAGKGLVIKHLMAGCTKVVVNRSEVRSAIREMPGIEIIVRPLDEMLACASEADVFTSTASETPLFLKEHVENLPOASPEVGLRHFVDISVPRNVGSCVGEVETARVYVDDLKEVVAANKEDRMKAMEAQTIITEESTQFEARNRDSLETPTTKLARAYAEIRIVASLEKCMKMGDDINKKTTTAVDDLSRGIVNRFHGMHRLRCGSDSRTLSLETLENHMLNRMYGLEKDLILEKLMAMAEQQQK"

1633..1663

/gene="At1g58290"

3'UTR

BASE COUNT 477 a 292 c 433 g 461 t

ORIGIN

Query Match

Best Local Similarity 35.3%; Score 679.6; DB 8; Length 1663;

Matches 985; Conservative 0; Mismatches 439; Indels 12; Gaps 3;

Qy 235 CCGTGGCAAGGCTGCCAGCGTCGCCCTCGAGCAGTTCAAGATAT---CCGCCGACC 291

Db 206 CTGGCTTAATGAGCTAGCATCTCTGCTCTTGAACAACTCAAGAACTGTCAGCTGATC 265

Qy 292 GGTACATGAAGAAAGAGTACCATAGCTGTGATAGGCTTCAGTGTACACACAGCACAG 351

Db 266 GATATCAAGAAAGAAAGAGCAGTATTGTTGTTGATTGGACTTAGTATTTCACACAGCTCTG 325

Qy 352 TGGAGATGCGTGAAGAACTTCTGTTGTCGAGAGAACTGTGGCCCGGCTGCTATTCAAGAAC 411

Db 326 TTGAGATGCGTGAAGAGCTTGTCTATACCTCAAGCTGAATGGCCCAAGAGCTATTGCTGAAT 385

Qy 412 TCACAGCTGACCATATTGAAGAGCGCTGCTCTTCTTAGTACCTGTAATAGAAATGAAA 471

Db 386 TGTGTGGTTTGAATCATATTAAGAAAGCTGTGTCTTAGTACTTGTAAACCGTATGGAGA 445
Qy 472 TTTATGTGGTGGCGCTATCATGGAACCGTGTGTATCAGAGAAAGTAGTGGACTGGATGTGCA 531
Db 446 TTTATGTGTAGCTTTATCTCAGCATCTGTGANGTAAAGAAAGTTACTGAAATGGATGCA 505
Qy 532 AGAAAAGTGTATTTCCCGCTTCCGAGCTTAGGAGACCTGTTCATCTTTGGCAAGAGTG 591
Db 506 AGACGAGTGGGATCCAGTTTTCAGAGATTTGTCAACACCGTTTCTGTGTGTACAAAGG 565
Qy 592 ATGCCACACGCTATCTGTTCAGAGTGTGAGTGGCTTGTGACTCTTTGGTTCTCCGTGAAG 651
Db 566 ATGCTACACAGATATATTTGAAGTCTCAGCTGGTCTTGACTCTCTTGTCTTAGGAGAG 625
Qy 652 GACAAATCCTTGTCTCAGGTTAAACAAAGTTGTGAGAGTGGACAGAAAGTGGAGGCTTGG 711
Db 626 GTCAGATCCTTGCAACAGGTGAACAAAGTTGTGAAAGTTGGTCAAGGAGTGAATGGCTTTG 685
Qy 712 GAAGAAACATCGATAGGATGTTCAAGATGCAATCAGTCTGCGAAAGCGTGTCCGACGG 771
Db 686 GGAGGAATATCAGCGGGCTGTTTAAACACGGGATAACTGTTGGTAAAGCGTGTAGAACAG 745
Qy 772 AGACCAACATATCATCTCGTGTCTTCTGTGAGTTCAGCGCGGTTGAACTGGGCCCTGA 831
Db 746 AGACAAATATTCCTTCTGTGGTGGTTCGGTTAGCTCAGCTGCTGTTGAACTTGTCTGCA 805
Qy 832 TGAAGCTTCCGAAGTCTGAAGCATGTGAGTGTAGATGCTTCTGATTCGTGTCTGTTAAA 891
Db 806 TGAAGCTTCCGAGTCTTCAACAGTATCTGTAGGATGTGTGTAATTCGGCGCTGGAAAA 865
Qy 892 TGGGAAGCTAGTGATCAACATCTGTTGCCAAAGGATGCAAGAGGTTGTGTGGTGA 951
Db 866 TGGGAAGCTTGTGATTAAGCATTTGATGGCTTAAGGGTTGCACAAAGTGTGTGTAGTCA 925
Qy 952 ACCGCTCCGTGAAAGGTTGATGCTATTCTGTGAGGATGAAGATATAGAGATCGTGT 1011
Db 926 ACAGAAGTGAAGAGAGGTTTCAGCTATCCGTGAGGAGATGCCGTGTTATGAGATTAT 985
Qy 1012 ACAGGCTCTCTCAGACATGTATCAAGCTGTGTGTAAGCTGATGTCTGTTCACCGCA 1071
Db 986 ATCGACCTCTTGATGAGATGCTAGCTTGTGTCTCAGAAAGCGGATGTGTGTATACGCA 1045
Qy 1072 CGCATCTGAACTTCTATTGTTGCAAAAGAACGACGACAGGCACTCCCGCTGTCTCTG 1131
Db 1046 CAGCTCTGAGCGCATTTGTTCTTAAAGGAGCATGTAGAGATCTCCCTCAAGCTTCTC 1105
Qy 1132 ATACTATGGAGGTTTCGCCCTGTTGTGACATATCTGTCCCGAGGAATGTCAAGCGAT 1191
Db 1106 CAGAGTTTGGAGGACTGAGGCAATTTTGTGACATCTCTGTTCCGAGAAATGTTGGATCTT 1165
Qy 1192 GTGTCTGTAAGTTGGCGCTGACAGTGTACAATGTTCGACGACTTGAAGAGGTTGGTGG 1251
Db 1166 GTGTTGGTGAAGTAGAACAAGCAGCGGGTTTCAATGTGGACGATCTCAAGAAAGTTGTTG 1225
Qy 1252 AAGCCAAACAGAGAGACCGGCTCAGAAAGCAATGGAGCGCAGACATCATCACCGAAG 1311
Db 1226 CTGCCAACAGAGACAGGATGAGAAAGCAATGGAAGCTCAAAACCAATATCACAGAG 1285
Qy 1312 AACTGAGACGTTTCAGGCAATGGAGGACTCGTGGAGACCGTTCCGACCATCAAGAAGC 1371
Db 1286 AATCCACACAGTTTGAAGCATGGAGGGATTCAATTGGAGACGGTTCCAAAGATCAAGAGT 1345
Qy 1372 TGAGTCTGTACCGGACAGGATCAGGCTCGGAGCTCGAAGTGCCTCGAGAAAGTAG 1431
Db 1346 TAAGAGCTTATGACAGAGAAATCAGAGTGGCAGAGCTCGAAGGTGCATGTCCAAAATGG 1405
Qy 1432 GTGAGGACCGCTCACCAAGAAAGATGAGGAGCCATCGAGGAGCTGAGACCGGCACTCG 1491
Db 1406 GAGATGA---TATCAACAAAGAAACACAGAGCGGTTGATGACTTAAGCCGAGGTATCG 1462
Qy 1492 TTAACAAGCTCTCCATCGCCGCTGACGACCTGAGGTGGCGGAGCGGACGACCGCGCA 1551

Db 1463 TGAACAGATTCTTACATGTTCCATGTCAGCATTTTGAGATGTGACGGAAGTGCACAGTAGAA 1522
Qy 1552 CCCTTTGACGAGACGCTCGAGAAACATGCACGCCCTCAACCGGATGTTTCAGGCTTCGACATGG 1611
Db 1523 CGCTGAGCGAGACCCCTTTGAGAAATATGCTCTTTGNAACAGAAATGTACGGTC-----TTG 1576
Qy 1612 AGAAGGCGATCATCGACAGAAAGATCAAGGCCAAGGTGGAGAACACAAAACCTGA 1667
Db 1577 AGAAAGACATTTTGGAGGAAAGCTTAAGGCAATGGCGGAACAACAACAGAAAGTAA 1632

Search completed: December 15, 2003, 14:57:28
Job time : 7206 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 12:46:12 ; Search time 4189 Seconds
(without alignments)
11163.001 Million cell updates/sec

Title: US-10-018-902-3
Perfect score: 1924
Sequence: 1 ccacgcgcgcgcacataa.....aaaaaaaaaaaaaaaaaaag 1924

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1923	99.9	1923	11	AY106615
c 2	829.4	43.1	831	10	BG319950
3	799.8	41.6	978	13	BQ619299
c 4	755	39.2	800	29	BZ748339

5	752	39.1	752	29	BZ748338
6	685.2	35.6	858	14	CB622460
7	673.8	35.0	854	29	CC424852
8	649.2	33.7	790	14	CB650182
9	632.6	32.9	799	14	CB672611
c 10	622.4	32.3	855	14	CB652061
11	621.4	32.3	812	14	CB621190
c 12	609.6	31.7	845	14	CB684065
13	606	31.5	751	14	CB621670
c 14	601.6	31.3	608	9	AW520155
15	584.2	30.4	733	14	CB685474
16	583.2	30.3	718	9	AU056848
17	583.2	30.3	749	13	BQ483617
c 18	570.6	29.7	786	14	CB670512
19	570.6	29.7	820	14	CB650183
c 20	563.2	29.3	781	14	CB623934
c 21	562	29.2	766	14	CB621191
22	555.8	28.9	559	14	CA452561
23	552.6	28.7	559	14	CA452567
c 24	550.6	28.6	813	14	CB669161
25	526	27.3	582	14	CD206259
26	518.8	27.0	679	13	BQ839101
c 27	508.8	26.4	512	12	B1993101
c 28	507.8	26.4	512	13	BQ539289
c 29	502.8	26.1	703	14	CB648319
c 30	489.6	25.4	543	9	AW671698
31	484	25.2	738	10	BP262195
c 32	481	25.0	719	12	BM084293
33	477.8	24.8	636	9	AV940881
34	476.4	24.8	600	12	BJ225193
35	471.6	24.5	579	10	BE599448
36	467.4	24.3	621	10	BE366884
c 37	467	24.3	478	12	BG842485
38	464.2	24.1	629	12	BJ452415
c 39	464	24.1	612	10	BE515566
c 40	463.4	24.1	621	14	CB859628
41	463.2	24.0	578	12	BJ214307
42	462.6	24.0	602	9	AV909911
43	461.8	24.0	645	12	B1952816
44	459.6	23.9	579	14	CD206237
45	457	23.8	580	10	BE594855

ALIGNMENTS

RESULT 1	AY106615	AV106615	1923 bp	mRNA	linear	HTC 16-OCT-2002									
LOCUS	Zea mays	PCO073362	mRNA sequence.												
DEFINITION	Zea mays	PCO073362	mRNA sequence.												
ACCESSION	AY106615														
VERSION	AY106615.1	GI:21209693													
KEYWORDS	HTC.														
SOURCE	Zea mays														
ORGANISM	Zea mays														
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.														
AUTHORS	Hayney, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.														
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes														
JOURNAL	Unpublished (2002)														
REFERENCE	2 (bases 1 to 1923)														
AUTHORS	Coe, E.H.														
TITLE	Direct Submission														
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA														
COMMENT	If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the														

BZ748338	PUDCK29TB
CB622460	OS11EA09D
CC424852	PURLJ39TD
CB650182	OSUNEB14H
CB672611	OSUNEE06L
CB652061	OSJNEC01H
CB621190	OS11EA06O
CB684065	OSJNEF13D
CB621670	OS11EA07L
AW520155	660061E06
CB685474	OSJNEF15O
AU056848	AU056848
BQ483617	WHE3510_F
CB670512	OSJNEE03J
CB650183	OSUNEB14H
CB623934	OS11EA12A
CB621191	OS11EA06O
CA452561	KRIN-2_F0
CA452567	KRIN-2_F0
CB669161	OSJNEE01H
CD206259	HS1_21_D0
BQ839101	WHE3589_D
B1993101	1020072H0
BQ539289	MEST614-D
CB648319	OSJNEB11K
AW671698	LG1_349_B
BP262195	HV_CBA000
BM084293	PPAP_05_C
AV940881	AV940881
BJ225193	BJ225193
BE599448	P11_87_F0
BE366884	P11_41_F0
BG842485	MEST31-D1
BJ452415	BJ452415
BE515566	WHE0614_B
CB859628	HI11A24W
BJ214307	BJ214307
AV909911	AV909911
B1952816	HVSWEM000
CD206237	HS1_21_D0
BE594855	P11_46_F0

maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES
source

Location/Qualifiers
1..1923
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:634031"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus Library"
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

BASE COUNT 504 a 465 c 550 g 404 t
ORIGIN

Query Match 99.9%; Score 1923; DB 11; Length 1923;
Best Local Similarity 100.0%; Pred. No. 6.7e-244;
Matches 1923; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCACCGCTCCGATCAATAAGAGGAGCTTGCGAAGTTGCCAAGGCTCCGATTTCCGCTA 60
DB 1 CCACCGCTCCGATCAATAAGAGGAGCTTGCGAAGTTGCCAAGGCTCCGATTTCCGCTA 60
QY 61 ATGCAGCATATATGCGACACAGCGTCAGCGACCCAGCGCGCGCGCAGCGCGCACCA 120
DB 61 ATGCAGCATATATGCGACACAGCGTCAGCGACCCAGCGCGCGCGCAGCGCGCACCA 120
QY 121 CCGCCAAAGCGCGGGGTGCTGCGGCTCTGCGAGAGGGTGGCGGGCGCGCGCACGC 180
DB 121 CCGCCAAAGCGCGGGGTGCTGCGGCTCTGCGAGAGGGTGGCGGGCGCGCGCACGC 180
QY 181 GCGCTCCGGGTGCTGCGGCTGCGACCGCGCGGTGCGAGCGCGCGCGCGCGCGCGTGG 240
DB 181 GCGCTCCGGGTGCTGCGGCTGCGACCGCGCGGTGCGAGCGCGCGCGCGCGCGTGG 240
QY 241 CAAGGCTCCAGCGCTCGCGCTCGAGCGATTCAGATATCGCGCGACCGGTACATGA 300
DB 241 CAAGGCTCCAGCGCTCGCGCTCGAGCGATTCAGATATCGCGCGACCGGTACATGA 300
QY 301 AGGAAGAGTACCATAGCTGTATAGGCTCTGAGGTACACAGCACCATGAGATGC 360
DB 301 AGGAAGAGTACCATAGCTGTATAGGCTCTGAGGTACACAGCACCATGAGATGC 360
QY 361 GTGAAAACTTGCTGTCAGAGGAACCTGTGCGCCCGTGTATTCAGAACTCACTAGCC 420
DB 361 GTGAAAACTTGCTGTCAGAGGAACCTGTGCGCCCGTGTATTCAGAACTCACTAGCC 420
QY 421 TGAACCATATTGAAGAGCTGCTGTATGATCTGTAATAGAAATTTATGTGG 480
DB 421 TGAACCATATTGAAGAGCTGCTGTATGATCTGTAATAGAAATTTATGTGG 480
QY 481 TGGCGCTATCATGGAACCGTGTATCAGAGAAGTAGTGCATGCGATGCGAAGAAAGTG 540
DB 481 TGGCGCTATCATGGAACCGTGTATCAGAGAAGTAGTGCATGCGATGCGAAGAAAGTG 540
QY 541 GTATTCGCGTTCGAGCTTAGGAGCACCTGTTTCATCTTCGGAAGCATGATGCCAC 600
DB 541 GTATTCGCGTTCGAGCTTAGGAGCACCTGTTTCATCTTCGGAAGCATGATGCCAC 600
QY 601 GCCATCTTTGAGGTGTCAGCTGGCTTGTATCTTTTGGTTCTCGGTGAAGCAAAATCC 660
DB 601 GCCATCTTTGAGGTGTCAGCTGGCTTGTATCTTTTGGTTCTCGGTGAAGCAAAATCC 660
QY 661 TTGCTCAGGTTAAACAAGTGTGAGGAGTGACAGAACAGTGGAGGCTTGGGAAGACA 720
DB 661 TTGCTCAGGTTAAACAAGTGTGAGGAGTGACAGAACAGTGGAGGCTTGGGAAGACA 720
QY 721 TCGATAGGATGTTCAAGGATGCAATCACTGCTGGAAGCGGTGTCGCGAGCGAGACCAACA 780

DB 721 TCGATAGGATGTTCAAGGATGCAATCACTGCTGGAAGCGGTGTCGCGAGCGAGACCAACA 780
QY 781 TATCATCTGGTGTCTTCTCTGTCAGTTCCAGCGGGTGAACCTGGCCCTGATGAAGCTTC 840
DB 781 TATCATCTGGTGTCTTCTCTGTCAGTTCCAGCGGGTGAACCTGGCCCTGATGAAGCTTC 840
QY 841 CGAAGTCTGAAGCACTGTCACTAGGATGCTTCTGATTTGGTGTGCTGTAATAATGGGAAGC 900
DB 841 CGAAGTCTGAAGCACTGTCACTAGGATGCTTCTGATTTGGTGTGCTGTAATAATGGGAAGC 900
QY 901 TAGTGATCAACATCTGTTTCCAAAGATGCAAGAGGTTGTTGTTGGTGAACCGCTCCG 960
DB 901 TAGTGATCAACATCTGTTTCCAAAGATGCAAGAGGTTGTTGTTGGTGAACCGCTCCG 960
QY 961 TGGAAAGGCTGCTATTCGTGAGGATGCTTCTGATTTGGTGTGCTGTAATAATGGGAAGC 1020
DB 961 TGGAAAGGCTGCTATTCGTGAGGATGCTTCTGATTTGGTGTGCTGTAATAATGGGAAGC 1020
QY 1021 TCTCAGACATGATCAAGCTGCTCTGAACTGATGCTGTTTCAACAGCACCGCATCTG 1080
DB 1021 TCTCAGACATGATCAAGCTGCTCTGAACTGATGCTGTTTCAACAGCACCGCATCTG 1080
QY 1081 AAACTTCAATTTTCCAAAGAAACACCGCAGGCACTCCCGCTGCTCTGATATATGG 1140
DB 1081 AAACTTCAATTTTCCAAAGAAACACCGCAGGCACTCCCGCTGCTCTGATATATGG 1140
QY 1141 GAGGTGTTGGCTGCTTGTGACATATCTGTCCTCCAGGAATGTCAGGSCATGTTGTCGTG 1200
DB 1141 GAGGTGTTGGCTGCTTGTGACATATCTGTCCTCCAGGAATGTCAGGSCATGTTGTCGTG 1200
QY 1201 AAGTTGGCGTTCGACAGTGTACAACTGTCGACGCTTGAAGAGGTTGGTGAAGCAACA 1260
DB 1201 AAGTTGGCGTTCGACAGTGTACAACTGTCGACGCTTGAAGAGGTTGGTGAAGCAACA 1260
QY 1261 AGGAGGACCGCTCAGGAAGCAATGAGGCGCAGACAACTCATCAACGAACTGAGAC 1320
DB 1261 AGGAGGACCGCTCAGGAAGCAATGAGGCGCAGACAACTCATCAACGAACTGAGAC 1320
QY 1321 GGTTCGAGGCAATGAGGAGGATCTGCTGAGGACCGTTCGACGACCATCAAGAACTGAGGTCGT 1380
DB 1321 GGTTCGAGGCAATGAGGAGGATCTGCTGAGGACCGTTCGACGACCATCAAGAACTGAGGTCGT 1380
QY 1381 ACGCGGACAGATCAGGCGCTCGAGCTCGAGAAAGTGCCTGCAAAAAGTAGGTGAGGACG 1440
DB 1381 ACGCGGACAGATCAGGCGCTCGAGCTCGAGAAAGTGCCTGCAAAAAGTAGGTGAGGACG 1440
QY 1441 CCCTCACCAGAAAGTACGAGGAGCCATCGAGGAGCTGAGCACCAGGATCGTTAACAGC 1500
DB 1441 CCCTCACCAGAAAGTACGAGGAGCCATCGAGGAGCTGAGCACCAGGATCGTTAACAGC 1500
QY 1501 TCCTTCATGCGCCCTCGAGCACCTGAGGTGCGACGCGCAGCGCAGCGCCCTTGAAGC 1560
DB 1501 TCCTTCATGCGCCCTCGAGCACCTGAGGTGCGACGCGCAGCGCAGCGCCCTTGAAGC 1560
QY 1561 AGAGCTCGAGAACATGCAAGCTTCAACCGGATGTTGAGCTCGACATGGAAGAGCGGA 1620
DB 1561 AGAGCTCGAGAACATGCAAGCTTCAACCGGATGTTGAGCTCGACATGGAAGAGCGGA 1620
QY 1621 TCATCGAGCAGAGATCAAGGCGGAGGAGACACAAACTGAGGCGCAGGAGCA 1680
DB 1621 TCATCGAGCAGAGATCAAGGCGGAGGAGACACAAACTGAGGCGCAGGAGCA 1680
QY 1681 TTTTCTTACCACTTATATATATATAGGCTTCCAATCTCATTTCCATTTTATTC 1740
DB 1681 TTTTCTTACCACTTATATATATATAGGCTTCCAATCTCATTTTATTC 1740
QY 1741 TTTCACTCAGTGAGCCCTTCCCTGCTCACTGTGATCGTTAACTGTTGTAATAGA 1800
DB 1741 TTTCACTCAGTGAGCCCTTCCCTGCTCACTGTGATCGTTAACTGTTGTAATAGA 1800
QY 1801 GCCATGCGAGGTTGTCAATAACAGCAATGTCCTCAATTTCCCAAGAGAGACT 1860


```

Db 1801 GCCATGGCAGCGTGTGTCAATTAACAGCAATGTGTCCCAATTCCTCCACAGAGAAGACT 1860
QY 1861 ATATTATATGATTTATTTGGAGCAAAATAGTTTACTTAAAAAAGAAAAAAGAAAAA 1920
Db 1861 ATATTATATGATTTATTTGGAGCAAAATAGTTTACTTAAAAAAGAAAAAAGAAAAA 1920
QY 1921 AAA 1923
Db 1921 AAA 1923

RESULT 2
BG319950/c
LOCUS
DEFINITION BG319950 831 bp mRNA linear EST 27-FEB-2001
cdna_clone Zm03_02b09 A Zm03 AAFEC ECORC cold stressed_maize seedlings Zea mays
ACCESSION BG319950
VERSION BG319950.1 GI:13149628
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS Simmonds,J.A., Singh,J.A., Piche,C., Cass,L., Couroux,P., De Moors
,A., Harris,L.J., Hattori,J.I., Ouellet,T., Robert,L.S., Sprout,D.
and Tinker,N.A.
TITLE Expressed Sequence Tags from Cold-Stressed Maize Seedlings Grown
Under High Light Intensity
JOURNAL Unpublished
COMMENT Contact: Singh,J.A.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-food Canada
KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, KIA
OC6, Canada
Tel: (613) 759-1662
Fax: (613) 759-1701
Email: singhja@em.agr.ca.
Location/Qualifiers
1. 831
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="CO328"
/db_xref="taxon:4577"
/clones="Zm03_02b09"
/tissue_type="Leaf, crown"
/dev_stage="4-leaf"
/clone_lib="Zm03 AAFEC ECORC cold stressed_maize seedlings"
/note="Vector: Bluescript SK+/XhoI-EcoRI; Site 1: Eco RI;
Site 2: Xho I; Corn seedlings at 4-leaf stage were exposed
to low temperature/high light (100C/700-800uE/m2/s) for 4
days. Plants were grown/treated by J. Simmonds/L. Cass.
Library prepared by C. Piche using Stratagene Kit."

BASE COUNT 172 a 224 c 219 g 216 t
ORIGIN

Query Match 43.1%; Score 829.4; DB 10; Length 831;
Best Local Similarity 99.9%; Pred. No. 5.7e-100;
Matches 830; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1009 TGTACAGCCCTCTCTCAGACATGATCAAGCTGCTGTAAGCTGATGTGTGTTCAACCA 1068
Db 831 TGTACAGCCCTCTCTCAGACATGATCAAGCTGCTGTAAGCTGATGTGTGTTCAACCA 772
QY 1069 GCACCGCATCTGAACCTTATTGTTCCGAAAGAACACGAGAGGCACTCCCCCTGTCT 1128
Db 771 GCACCGCATCTGAACCTTATTGTTCCGAAAGAACACGAGAGGCACTCCCCCTGTCT 712
QY 1129 CTGATACTATGGGAGGTGTTGCGCTGTTGTGCGACATATCTGTCCTCCAGGAATGTACGG 1188
Db 711 CTGATACTATGGGAGGTGTTGCGCTGTTGTGCGACATATCTGTCCTCCAGGAATGTACGG 652

```

```

QY 1189 CATGTGTCTCTGAAGTTTGGCCCTGTCACGAGTGTCAATGTTCGACGACTTGAAGAGGTGG 1248
Db 651 CATGTGTCTCTGAAGTTTGGCCCTGTCACGAGTGTCAATGTTCGACGACTTGAAGAGGTGG 592
QY 1249 TGGAAAGCCAAACAGAGGAGCAGCGGCTCAGGAAAGCAATGAGGCGCAGCAATCATCACCG 1308
Db 591 TGGAAAGCCAAACAGAGGAGCAGCGGCTCAGGAAAGCAATGAGGCGCAGCAATCATCACCG 532
QY 1309 AAGAACTGAGAGCGGTTGAGGCGATGGAGGAGCTGCTGGAGACCGTTCCGACCATCAAGA 1368
Db 531 AAGAACTGAGAGCGGTTGAGGCGATGGAGGAGCTGCTGGAGACCGTTCCGACCATCAAGA 472
QY 1369 AGCTGAGGTCCGTAGCGGACAGGATCAGGGGCTCGAGACTCGAGAAGTGCCTGCGAGAAG 1428
Db 471 AGCTGAGGTCCGTAGCGGACAGGATCAGGGGCTCGAGACTCGAGAAGTGCCTGCGAGAAG 412
QY 1429 TAGGTGAGGACGCCCTCACCAAGAGATGAGGAGAGCCATCGAGGAGCTGAGACACCGCA 1488
Db 411 TAGGTGAGGACGCCCTCACCAAGAGATGAGGAGAGCCATCGAGGAGCTGAGACACCGCA 352
QY 1489 TCGTTACAGCTCCTCCATGCGCCGCTGAGGACCTGAGGTGCGACGCGGACGACGCC 1548
Db 351 TCGTTACAGCTCCTCCATGCGCCGCTGAGGACCTGAGGTGCGACGCGGACGACGCC 292
QY 1549 GCACCTTGGACGAGCGCTCGAGAAACATGCAACGCGCTCAACCGGATGTTCAAGCCTCGACA 1608
Db 291 GCACCTTGGACGAGCGCTCGAGAAACATGCAACGCGCTCAACCGGATGTTCAAGCCTCGACA 232
QY 1609 TGGAGAGGCGGATCATCGAGCAGAGATCAAGGCCAAGGTGGAGAAGACACAAAACCTGAG 1668
Db 231 TGGAGAGGCGGATCATCGAGCAGAGATCAAGGCCAAGGTGGAGAAGACACAAAACCTGAG 172
QY 1669 GCCAGGAAGCAATTTTCTACCAACATATCTATATATAGCGTCTCCAAATCTCATTTCC 1728
Db 171 GCCAGGAAGCAATTTTCTACCAACATATCTATATATAGCGTCTCCAAATCTCATTTCC 112
QY 1729 ATTTTATTTATCTTCTACTCAGTCAGCGCCCTTCCCTGCTCACTGTGATCGTTAACTGCT 1788
Db 111 ATTTTATTTATCTTCTACTCAGTCAGCGCCCTTCCCTGCTCACTGTGATCGTTAACTGCT 52
QY 1789 CTGTGAATTAGAGCCATGGCAGCGTGTGTCAATTAACAGCAATGTGTCCCA 1839
Db 51 CTGTGAATTAGAGCCATGGCAGCGTGTGTCAATTAACAGCAATGTGTCCCA 1

RESULT 3
BG619299
LOCUS
DEFINITION RNOSE06A12_SK.ab1 Salt stressed Zea mays roots cDNA library Zea
mays cDNA Clone RNOSE06A12_SK.ab1 similar to (AC078840) putative
glutamyl-tRNA reductase [Oryza sativa], mRNA sequence.
ACCESSION BG619299
VERSION BG619299.1 GI:21621293
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
AUTHORS Wang,H. and Bohnert,H.J.
TITLE Genomics of plant stress tolerance
JOURNAL Unpublished
COMMENT Contact: Mark Fredricksen
Department of Plant Biology
University of Illinois
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 2172655473
Email: Dohnertlab@life.uiuc.edu.
Location/Qualifiers
1. .978
/organism="Zea mays"
/mol_type="mRNA"

FEATURES
source

```

Db 854 CCCGCGATTCGGAACCTCTCATTTGTTGCGAAAGAACACCCCGAGGCGACCTTTCC 910

RESULT 4
BZ748339/c

LOCUS BZ748339 800 bp DNA linear GSS 03-MAR-2003
DEFINITION PUDCK29TD ZM 0.6_1.0_KB Zea mays genomic clone ZMMBTa155F09,
genomic survey sequence.

ACCESSION BZ748339
VERSION BZ748339.1 GI:28728049
KEYWORDS GSS.

SOURCE Zea mays
ORGANISM Zea mays

REFERENCE 1 (bases 1 to 800)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick
A., Frazer,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.

TITLE Maize Genomics Consortium
COMMENT Unpublished
Other_GSSs: PUDCK29TB
Contact: Cathy Whitelaw
TIGR

7912 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.

FEATURES
source 1..800
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTa155F09"
/notes="Vector: pCR4-TOPO; Site 1: EcoRI, 0.6-1.0 kb high
cot selected genomic DNA library"
BASE COUNT 174 a 214 c 200 g 212 t

Query Match 39.2%; Score 755; DB 29; Length 800;
Best Local Similarity 99.1%; Pred. No. 3.5e-90;
Matches 780; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

Qy 1114 CACTCCGCCCTGCTCTGATACATATGGAGGTGTTCCCTCTTTTGTGCAATATCTGTCC 1173
Db 799 CTCGCCCTCTGTTCTCTGATCTATGCGAGGTGTTCCCTCTTTTGTGCAATATCTGTCC 740

Qy 1174 CCA-GGAATGTCAAG-CCGATGTGTCTGAAGTTGGCGCTGCACAGTGTACAAATCTCGA 1231
Db 739 CCAAGGAATGTCAAGCGCATGTGTGTGAAGTTGGCGCTGCACAGTGTACAAATCTCGA 680

Qy 1232 CGACTTTGAAAGAGGTGGTTCGAAGCCCAACGAAGGAGGACCCGGCTCAGGAAGCAATGAGGC 1291
Db 679 CGACTTTGAAAGAGGTGGTTCGAAGCCCAACGAAGGAGGACCCGGCTCAGGAAGCAATGAGGC 620

Qy 1292 GCAGACAATCATCACCGAAGACTGAGACGGTTCCGAGGCATGGAGGACTCCGCTCGAGAC 1351
Db 619 GCAGACAATCATCACCGAAGACTGAGACGGTTCCGAGGCATGGAGGACTCCGCTCGAGAC 560

Qy 1352 CGTTCGAGCATCAAGAAGCTGAGGTCTGATCGCGACAGCAATCAGGCCCTCGAGACTCGA 1411
Db 559 CGTTCGAGCATCAAGAAGCTGAGGTCTGATCGCGACAGCAATCAGGCCCTCGAGACTCGA 500

Qy 1412 GAAGTCCCTGCAGAAAGTAGGTGAGGACGCCCTTCACCAAGAAAGATGAGGAGGCCATCGA 1471
Db 499 GAAGTCCCTGCAGAAAGTAGGTGAGGACGCCCTTCACCAAGAAAGATGAGGAGGCCATCGA 440

Qy 1472 GGAGCTTGAGCACCGGCATCGTTAAACAAGCTCTCTCAATGCGCCGCTGACGACCTCAGGTG 1531

Db 439 GGAGCTGAGCACCGGCATCGTTAAACAAGCTCTCCATGGCCCGCTGCAGCACCTGAGGTG 380

QY 1532 CGAGCGGACGACAGCGCCACCCCTTGACGAGACGCTCGAGAACATGACGCGCCCTCAACCG 1591

Db 379 CGAGCGGACGACAGCGCCACCCCTTGACGAGACGCTCGAGAACATGACGCGCCCTCAACCG 320

QY 1592 GATGTTTCAGCCTCGACATCGAGAGAGCGGATCATCGAGCAGAGATCAAGGCGCAAGGTGGA 1651

Db 319 GATGTTTCAGCCTCGACATCGAGAGAGCGGATCATCGAGCAGAGATCAAGGCGCAAGGTGGA 260

QY 1652 GAAGACACAAACTGAGCGCCAGGAGCAATTTTCTACCAACATATCTATATATATAGC 1711

Db 259 GAAGACACAAACTGAGCGCCAGGAGCAATTTTCTACCAACATATCTATATATATAGC 200

QY 1712 GTCTCCCAATCTCAATTCATTTTATCTCTTTCATCTAGTCAGCGCCCTTCCCTGCTCACT 1771

Db 199 GTCTCCCAATCTCAATTCATTTTATCTCTTTCATCTAGTCAGCGCCCTTCCCTGCTCACT 140

QY 1772 GTGATCGTTAACTGTGCTGTAATTAGAGCCATGAGCGAGCGTGTGTCAATTAACAGCAAT 1831

Db 139 GTGATCGTTAACTGTGCTGTAATTAGAGCCATGAGCGAGCGTGTGTCAATTAACAGCAAT 80

QY 1832 GTGTCCCAATTTCCCAACAGAGAAAGACTATATTTATATGTCATTTATTTGGAGCAATAGT 1891

Db 79 GTGTCCCAATTTCCCAACAGAGAAAGACTATATTTATATGTCATTTATTTGGAGCAATAGT 20

QY 1892 TTACTTA 1898

Db 19 TTACTTA 13

RESULT 5
B2748338
LOCUS
DEFINITION PUDCK29TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMB155F09,
genomic survey sequence.

ACCESSION B2748338
VERSION B2748338.1 GI:28728048
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
AUTHORS 1 (bases 1 to 752)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick
A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.
Maize Genomics Consortium
Unpublished
Other_GSSs: PUDCK29TD
Contact: Cathy Whitelaw
TIGR
712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source
1..752
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMB155F09"
/notes="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
cot selected genomic DNA library"
BASE COUNT 198 a 183 c 230 g 141 t

Query Match 39.1%; Score 752; DB 29; Length 752;
Best Local Similarity 100.0%; Pred. No. 8.9e-90;

Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 879 GGTCTCGTAAATATGGGAAGCTAGTATCAAAATCTGTTGTCCTCAAGGATGCAAG 938

Db 1 GGTCTCGTAAATATGGGAAGCTAGTATCAAAATCTGTTGTCCTCAAGGATGCAAG 60

QY 939 GTTCTGTGGTGAACCGCTCGTGGAAAGGATGATCTATTCGTGAGAGATGAAGAT 998

Db 61 GTTCTGTGGTGAACCGCTCGTGGAAAGGATGATCTATTCGTGAGAGATGAAGAT 120

QY 999 ATAGAGATCTGTATACGGCTCTCTCAGACATGATCAAGCTGCTGTAAGCTGAT 1058

Db 121 ATAGAGATCTGTATACGGCTCTCTCAGACATGATCAAGCTGCTGTAAGCTGAT 180

QY 1059 GTGTTTCAACGACCGCATCTGAAATCTTATTTGTCGAAAGAACCGCAGAGCACTC 1118

Db 181 GTGTTTCAACGACCGCATCTGAAATCTTATTTGTCGAAAGAACCGCAGAGCACTC 240

QY 1119 CCCCTGTCTCTGATATCTATGGGAGGTGTTGCGCCTGTTTGTGACATATCTGTCCCAGG 1178

Db 241 CCCCTGTCTCTGATATCTATGGGAGGTGTTGCGCCTGTTTGTGACATATCTGTCCCAGG 300

QY 1179 AATGTCAGCGCATGTGTCTGAAGTTGGCGCTGCACAGTGTACAATGTCGACACTTG 1238

Db 301 AATGTCAGCGCATGTGTCTGAAGTTGGCGCTGCACAGTGTACAATGTCGACACTTG 360

QY 1239 AAAGAGTGTGGAGCAACAGAGGAGCGCGCTCAGGAAAGCAATGGAGCGCAGACA 1298

Db 361 AAAGAGTGTGGAGCAACAGAGGAGCGCGCTCAGGAAAGCAATGGAGCGCAGACA 420

QY 1299 ATCATCACGAAGAACTGAGACGCTCGAGCGCTCGAGGAGGACTCGCTGGAGACGTTCCG 1358

Db 421 ATCATCACGAAGAACTGAGACGCTCGAGCGCTCGAGGAGGACTCGCTGGAGACGTTCCG 480

QY 1359 ACCATCAAGAAGCTGAGTGTGTAGCGGACAGGATCAGGGCTCGAGAGCTCGAGAAGTGC 1418

Db 481 ACCATCAAGAAGCTGAGTGTGTAGCGGACAGGATCAGGGCTCGAGAGCTCGAGAAGTGC 540

QY 1419 CTGCGAAGAGTGTGAGGAGCGCTCACCAGAGAGATGAGGAGAGCCATCGAGAGCTG 1478

Db 541 CTGCGAAGAGTGTGAGGAGCGCTCACCAGAGAGATGAGGAGAGCCATCGAGAGCTG 600

QY 1479 AGCACCGGCATCGTTAAACAAGCTCTCCATGGCCCGCTGCAGACCTGAGGTGCGACGGC 1538

Db 601 AGCACCGGCATCGTTAAACAAGCTCTCCATGGCCCGCTGCAGACCTGAGGTGCGACGGC 660

QY 1539 AGCGACAGCGCACCCCTTGACGAGAGCTCGAGAACATGCAACCGCTCAACCGGATGTC 1598

Db 661 AGCGACAGCGCACCCCTTGACGAGAGCTCGAGAACATGCAACCGCTCAACCGGATGTC 720

QY 1599 AGCTCGACATGAGAGAGCGCATCTCGAGCA 1630

Db 721 AGCTCGACATGAGAGAGCGCATCTCGAGCA 752

RESULT 6
CB622460
LOCUS
DEFINITION OSIIIEa09D05.f OSIIIEa Oryza sativa (indica cultivar-group) cDNA
clone OSIIIEa09D05 5', mRNA sequence.

ACCESSION CB622460
VERSION CB622460.1 GI:29617448
KEYWORDS EST.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartioideae; Oryzae; Oryza.
REFERENCE 1 (bases 1 to 858)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea


```

QY 837 CTTCCGAAGCTGAAGCACTGTACAGTGGTCTTCTGATTGGTGGTAAATGGGA 896
DB 601 CTTCCAAAGTCCGAATGCTATCTCTAGGATGCTGTTGATTGGTGGCAAGATGGGA 660
QY 897 AAGCTAGTAGTCAAAACATCTGTTCCCAAGGATGCAAGAGGTTGTTGGTGAACCGC 956
DB 661 AAGTTGGTGGTAAACATTTGATTGCCAAGGATGCAAGAAAGTTGTTGGTGAACCGT 720
QY 957 TCCGTGGAAAGGTTGGATGCTATTCTGTGAGGAGATGAAGATATAGAGATCGTGTACAGG 1016
DB 721 TCAGTGGAAAGGTTGGATGCTATCGCGAAGAGATGAAGACATTGAGATTGTGTACAGG 780
QY 1017 CCTCTCTCAG 1025
DB 781 CCTCTTACAG 790

RESULT 9
LOCUS CB672611
DEFINITION OSJNE06L17.7 f OSJNEe Oryza sativa (japonica cultivar-group) cDNA
ACCESSION CB672611
VERSION CB672611.1 GI:29676336
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 799)
Jantasuriyarat, C., Lu, G., Gowda, M., Hatfield, J., Zhou, B., Mazur, E.,
Kudrna, D., Dean, R., Soderlund, C., Wing, R. and Wang, G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 06 row: L column: 17
Seq primer: gta aaa cga cgg cca gtc.
Location/Qualifiers
1..799
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNE06L17"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEe"
/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 24 hrs after inoculation with Rice Blast (70-15)"

BASE COUNT 202 a 156 c 229 g 212 t
ORIGIN

Query Match 32.9%; Score 632.6; DB 14; Length 799;
Best Local Similarity 87.0%; Pred. No. 4.2e-74;
Matches 695; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 444 GTTCTTACTACTGTAATAGATGCAATTTATGCTGGCGCTATCATGGAACCGTGT 503
DB 1 GTTCTTAGCTGCTTATAATGGAATCTATGTGTAGTCTTTATCTGTGAACCGTGG 60

```

```

QY 504 ATCAGAGAGTAGTGGACTGATGTGGAAGAAAAGTGTGTATTCGGCTTCCGAGCTTAGG 563
DB 61 ATTAGAGAAGTGGTAGACTGGATGTCAAAGAAAAGTGGAAATCCCTGCTTTCGAGCTCAGG 120
QY 564 GAGCACCTGTTTCATCTTCGGAAGCAGTGTATGCCACACGCCATCTGTTTGGAGGTGTGAGCT 623
DB 121 GAGCATCTATTTCATGTTGCGTGCAGTGTATGCCACACGCCATCTGTTTGGAGGTATCTGCT 180
QY 624 GGCCTTGATCTCTTTGGTTCTCGGTGAAGGACAAATCTCTGTGCTCAGGTTAAACAAGTTGTG 683
DB 181 GGGCTTGACTCTTTGGTTCTTTGGAGAGGGCAATCTCTGCTCAAGTTTAAACAAGTTGTGTC 240
QY 684 AGGAGTGGACAGAACAGTGGAGGCTTGGGAAAAGAACATCGATAGATGATTTCAAGATGCA 743
DB 241 AGAAGTGGGCAAAAACAGTGGAGGCTTGGGAAAAGAACATCGATAGATGATTTCAAGATGCA 300
QY 744 ATCACTGCTGAAAAGCGTGTCCGAGCGAGAGCAACATATCATCTGGTGGTCTGTTCTGTCTC 803
DB 301 ATCACTGCTGAAAAGCGTGTCCGTCGAGAGACTAACATATCATCAGGTGCTGTCTGTCTC 360
QY 804 AGTTGAGCGGCGGTGAACCTGCGCCTGATGAAGCTTCCGAAGTCTGAAGCACTGTGAGCT 863
DB 361 AGTTGAGCTGCAGTTGAATTTGGCCTTGATGAAGCTTCCAAAGTCCGAATGCTATCTGCT 420
QY 864 AGGATGCTTCTGATTTGGTGTGTTAAATGGGAAAGCTAGTATCAACATCTGGTTGCC 923
DB 421 AGGATGCTTGTGATTTGGTGTGTTGCAAGATGGGAAAGTTGGTGTGTTAAACATTTGATGCC 480
QY 924 AAAGATGCAAGAAGTTGTTGTTGTTGTAACCGCTCCGTGGAAAAGGTTGGATGCTATTGCT 983
DB 481 AAGGATGCAAGAAGTTGTTGTTGTTGTAACCGCTTCAGTGGAAAAGGTTGGATGCTATCGC 540
QY 984 GAGGAGATGAAGAATATAGAGATCGTGTACAGGCTCTCTCAGACATCATCAAGCTGCT 1043
DB 541 GAAGAGATGAAGAATATGAGATTTGTTGTACAGGCTCTCTTACAGAGATGTTATGAAGCCGT 600
QY 1044 GCTGAAGCTGATGCTGTTTCCACGACCGCATCTGAACTTCTGTTTTCGCAAGAA 1103
DB 601 GCCGAAGCTGATGTTGTTGTTTCCAAAGCAAGCGCATCCGAAACCCCATTTGTTCAAAAGAG 660
QY 1104 CACGACAGGCACTCCCCCTGCTCTGATATACTATGGAGGTTGTCGCTGTTTCTCGAC 1163
DB 661 CACGACAGGCGCTTCCGCTATTCTGATGCTATGTTGGTGTGTTGCTGCTCTTTGTCAC 720
QY 1164 ATATCTGTCCCAGGAATGTTCAGGCGCATGTGTGCTGAAGTTGGCGTGCAGAGTGTAC 1223
DB 721 ATATCCGTCGCCAGAAATGTTCAGGCGCTGTGTGTTGAAAGTTGGCCATGCGCGAGTATAC 780
QY 1224 AATCTCGACGACTTGAAG 1242
DB 781 AACGTCGATGACTTGAAG 799

RESULT 10
LOCUS CB652061/c
DEFINITION OSJNEC01H23.r OSJNEc Oryza sativa (japonica cultivar-group) cDNA
ACCESSION CB652061
VERSION CB652061
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 855)
Jantasuriyarat, C., Lu, G., Gowda, M., Hatfield, J., Zhou, B., Mazur, E.,
Kudrna, D., Dean, R., Soderlund, C., Wing, R. and Wang, G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished

```


COMMENT

Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: <http://genome.arizona.edu>
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 01 row: H column: 23
Seq primer: gga aac agc tat gac cat g
Location/Qualifiers
1. .855
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEC01H23"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEC"
/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"
BASE COUNT 186 a 231 c 205 g 233 t
ORIGIN

FEATURES

source

Query Match 32.3%; Score 622.4; DB 14; Length 855;
Best Local Similarity 85.6%; Pred. No. 8.9e-73;
Matches 692; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
Qy 912 CATCTGTTGCAAGGATGCAAGAGTGTCTGCTGTAACCGCTCGGTGGAAAGGTTG 971
Db 855 CATTTGATGCCAAGGATGCAAGAAAGTTGTTGTTGTAACCGCTTCAGTGGAAAGGTTG 796
Qy 972 GATGCTATTCTGAGGAGATGAAAGATATAGAGATCGTGTACAGCGCTCTCTCAGACATG 1031
Db 795 GATGCCATCCGGAAGAGATGAAAGACATGAGATTGTGTACAGGCTCTTACAGAGATG 736
Qy 1032 TATCAAGCTGCTGTAAGCTGATGTGTTGTACACGACCGCATCTGAAACTTCATTG 1091
Db 735 TATGAAGCCGCTGCGAAGCTGATGTGTTGTTCACAGCACGGCATCCGAAACCCCATTTG 676
Qy 1092 TTCCAAAGAACACGACAGAGGCACTCCCGCTGCTCTGATCTACTATGGGAGGTTCGC 1151
Db 675 TTCAAAAGAGACACGACAGAGGCGCTTCCCGCTATTCTGATGCTATGGGTGTTTCGA 616
Qy 1152 CTGTTTGTGACATATCTGTCCCAGGAATGTGTCAGCGCATGTGTCTGAAAGTTGGCGCT 1211
Db 615 CTCTTTGTGACATATCCGTCCCCAGAAATGTGTCAGCGCTGTGTCTGAAGTTGGCCAT 556
Qy 1212 GCAGGAGTGTACATGTGAGAGCTTGAAGAGTGTGGAAGCCAAAGGAGGACCGG 1271
Db 555 GCOCGAGTATACAACTGTCATGCTTGAAGAGAGTGTGGAAGCCAAAGGAGGACCGG 496
Qy 1272 CTCAGGAAAGCAATGGAGGCGGACACATCATCACCGAAGCACTGAGCGGTTTCAGGCA 1331
Db 495 CTTAGGAAAGCAATGGAGGCGGACCAATCATCACCAAGAAATGAAAGCGTTTCAGGCA 436
Qy 1332 TGGAGGAGCTCGTGGAGACCGTTTCCGACCATCAAGAGCTGAGGTGCTACGCGGACAGG 1391
Db 435 TGGAGGAGCTCGTGGAGACTGTTCCGACTATCAAGAGCTGAGGTCTACGCGGACAGG 376
Qy 1392 ATCAGGCGCTCGAGCTCGAAGAGTGCCTCGAAGAGTAGGTGAGGACGCGCTTCAACAAG 1451
Db 375 ATCAGGCGCTCGAGCTTGAAAGAGTGCCTCAGAAGATCGGCGCAAGACGCGCTTCAACAAG 316
Qy 1452 AAGATGAGGAGAGCCATCGAGGAGCTGAGCACCAGGCTCGTTTAAAGCTCTCTCCATGCG 1511
Db 315 AAGATGAGAGATCCATCGAGGAGCTGAGCACCAGGCTCGTTTAAAGCTCTCTCCACGGC 256

Qy 1512 CCGCTGCAGCACCTGAGTGGAGCGGACGACGACGCGCTTTCAGACGCGTGGAG 1571
Db 255 CCATTGCGAGCACCTGAGATGTGACGCGAGCGACGCGCCCTTCGATGACGCGTGGAG 196
Qy 1572 AACATGCAACCCCTCAACCGGATGTTTCAGCTTCGACATGGAGAGGCGATCATCGAGCAG 1631
Db 195 AACATGCAACCCCTCAACAGGATGTTTCAGCTTCGACACCGGAGAGGCGATCATTTGAGCAG 136
Qy 1632 AAGATCAAGCCCAAGTGGAGAGACACAAACTGAGCCAGGAGCAATTTTCTTACCA 1691
Db 135 AAGATCAAGCCCAAGTGGAGAGATGCCAGAACTGAGATTGAAGAGAGATTTTTTTTTT 76
Qy 1692 CCATTATCTATATATATAGGCTCTCCAA 1719
Db 75 TCAGCCCGTGTACTTACTATGTATACTA 48

RESULT 11

CB621190
LOCUS CB621190
DEFINITION Oryza sativa (indica cultivar-group) cDNA
clone OSIIa06012.5', mRNA sequence.
CB621190
ACCESSION CB621190.1 GI:29616178
VERSION
KEYWORDS EST.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 812)
Jantauriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: <http://genome.arizona.edu>
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 06 row: O column: 12
Seq primer: gta aaa cga cgg cca gtc.
Location/Qualifiers
1. .812
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="mRNA"
/cultivar="IR36"
/db_xref="taxon:39946"
/clone="OSIIa06012"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSIIa"
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; Lesion Mimic SPL 11"
BASE COUNT 192 a 171 c 258 g 190 t
ORIGIN

FEATURES

source

Query Match 32.3%; Score 621.4; DB 14; Length 812;
Best Local Similarity 87.4%; Pred. No. 1.2e-72;
Matches 692; Conservative 0; Mismatches 97; Indels 3; Gaps 1;
Qy 165 GCCGCGCGGACGCGCGCTCCCGGGTGGTGGCGTGGCGAGCCCGCGCGTGG---AG 221
Db 21 GCCGCGCGGACGCGCGCTCCCGCGTGGTGGTGGCGAGCCCGCGCGTGGAGGCGCAG 80

Query Match

Best Local Similarity 87.4%; Score 621.4; DB 14; Length 812;
Matches 692; Conservative 0; Mismatches 97; Indels 3; Gaps 1;
Qy 165 GCCGCGCGGACGCGCGCTCCCGGGTGGTGGCGTGGCGAGCCCGCGCGTGG---AG 221
Db 21 GCCGCGCGGACGCGCGCTCCCGCGTGGTGGTGGCGAGCCCGCGCGTGGAGGCGCAG 80

[illegible]

RESULT 12	CB684065	845 bp	mRNA	linear	EST 09-APR-2003
LOCUS	OSJNefl3D18_r	OSJNefl3D18_r	Oriza sativa	(japonica cultivar-group)	cDNA
DEFINITION	clone OSJNefl3D18 3', mRNA sequence.				
ACCESSION	CB684065				
VERSION	CB684065.1	GI:29687790			
KEYWORDS	EST.				
SOURCE	Oriza sativa (japonica cultivar-group)				
ORGANISM	Oriza sativa (japonica cultivar-group)				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.				
REFERENCE	1 (bases 1 to 845)				
AUTHORS	Jantasuriyarat, C., Lu, G., Gowda, M., Hatfield, J., Zhou, B., Mazur, E., Kudrna, D., Dean, R., Soderlund, C., Wing, R. and Wang, G.				
TITLE	Large-scale identification of ESTs involved in the interaction between rice and Magnaporthe grisea				
JOURNAL	Unpublished				
COMMENT	Contact: Rod Wing				

Arizona Genomics Institute
 University of Arizona
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
 85721-0088, USA
 Tel: 520 626 3967
 Fax: 520 621 9288
 Email: <http://genome.arizona.edu>
 PCR Primers
 FORWARD: gta aaa cga cgg cca gtg
 BACKWARD: gga aac agc tat gac cat g
 Plate: 13 row: D column: 18
 Seq primer: gga aac agc tat gac cat g.
 Location/Qualifiers
 1..845
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="OSJNFI13D18"
 /tissue_type="Leaf"
 /dev_stage="3 week"
 /lab_host="DH10B"
 /clone_lib="OSJNEf"
 /notes="vector: pBluescript II KS +; Site_1: EcoRI; Site_2: XhoI; Uninfected Control"
 BASE COUNT 179 a 228 c 200 g 238 t
 ORIGIN

Query Match 31.7%; Score 609.6; DB 14; Length 845;
 Best Local Similarity 85.6%; Pred. No. 4.3e-71;
 Matches 678; Conservative 0; Mismatches 114; Indels 0; Gaps 0

QY 928 GATCAAGAAGGTTGTGTGTCGAACCGCTCCGTGGAAAGGGTGGATGCTATTTCGTGAGG 987
 Db 845 GATCAAGAAGTTGTGTGGTAACCGTTTCAGTGAAGGGTGGATGCCATCCGGAAG 786

QY 988 AGATGAAGATATAGAGATCGTTACAGGCCTCTCTCAGACATGTATCAAGCTGCTGCTG 1047
 Db 785 AGATGAAGACATTGAGATTGTGTACAGGCCTCTTACAGAGATGTATGAAGCCGTGCCG 726

QY 1048 AAGCTGATGTCTGTTTCACGACCGCGATCTGAAACTTCATTTGCGAAAAAGAACAG 1107
 Db 725 AAGCTGATGTCTGTTTCACAGCAGCGCATCCGAAACCCCATTTGTTCAAAAGGAGCAG 666

QY 1108 CAGAGGCATCCCCCTCTCTGTACTATGGAGGTGTTCGCCCTGTTTGTTCGACATAT 1167
 Db 665 CAGAGGCCTCCCGCTATTTCTGTATGCTATGGTGTGTTCGACTCTTTGTTCGACATAT 606

QY 1168 CTGTCCCCAGGAATGTCAAGCATGTGTGTCTGAACTTGGCGTGCACGAGTGTACAATG 1227
 Db 605 CCGTCCCCAGAAATGTCAGGCCCTGTGTCTGAACTTGGCCATGCGCGAGTATACAAG 546

QY 1228 TCGACGACTTGAAGAGGTGGTGGAGCCAAACAGGAGGACCGGCTCAGGAAAGCAATGG 1287
 Db 545 TCGATGACTTGAAGAGGTTGTGGAAAGCCAAACAGGAGGACCGGCTTAGGAAAGCAATGG 486

QY 1288 AGCGCAGACAATCATCATCCGAAGACTCAGACGGTTTCAGGCATCGAGGGACTCGCTCG 1347
 Db 485 AGGCCCAACAAATCATCCCAAGAAATTGAAACGGTTTCAGGCATCGAGGGACTCGCTCG 426

QY 1348 AGACCGTTCGACCATCAAGAAGCTGAGTTCGTACCGCGACAGGATCAGGGCTCCGAGC 1407
 Db 425 AGACTGTTCCGACTATCAAGAAGCTGAGTTCCTACGCCACAGGATCAGGGCTTCGAGC 366

QY 1408 TCGAAGTGTCCTCAGAAAGTAGGTGAGGACGCCCTCAACGAAGAAGATGAGGAGGCCA 1467
 Db 365 TTGAGAAGTGTCCTCAGAAAGATCGGCGAAGAGCGCCCTCAACGAAGAAGATGAGAAGATCCA 306

QY 1468 TCAGAGAGCTGAGCACCGGCATCGTTAAACAGCTCCTCCATGGCCCGCTGCAGCACCTGA 1527
 Db 305 TCAGAGAGCTCAGCACCGGCATCGTGAACAGCTTCTCCACGGCCCATTTGCAGCACCTGA 246

QY 1528 GGTGCGACGCGCAGCAGCCGACCCCTTTGACGAGACGCTTCAGAAACATGCACGCCCTCA 1587

```

Db      245  GATGTGACGGCAGCAGCCGCTCGATGAGACGCTGAGAACATGCGACGCCCTCA 186
QY      1588  ACCGATGTTTCAGCTTCGACATGGAGAGCCATCATCGAGCAGAGATCAAGGCCAAGG 1647
Db      185  ACAGGATGTTACGCTTCGACACCGAGAGAGCCATCATTCAGCAGAGATCAAGGCCAAGG 126
QY      1648  TGGAGAAGACACAAACTGAGGCCAGGAGCAATTTTCTACCAACCATTTATCTATATATA 1707
Db      125  TGGAGAAGTCCAGAACTGAGATTGAAGAAGAGATTTTTTTTTTTCAGCCCGTGTATCTA 66
QY      1708  TAGCTCTCCAA 1719
Db      65  CTATGTATACTA 54

```

RESULT 13

```

CB621670
LOCUS      751 bp mRNA linear EST 08-APR-2003
DEFINITION OSIIIEa07L19.f OSIIIEa Oryza sativa (indica cultivar-group) cDNA
SOURCE      clone OSIIIEa07L19 5', mRNA sequence.
ACCESSION   CB621670
VERSION     CB621670.1 GI:29616658
KEYWORDS
SOURCE
ORGANISM

```

```

Oryza sativa (indica cultivar-group)
Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 751)
Jantasuriyatat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tet gac cat g
Plate: 07 row: L column: 19
Seq primer: gta aaa cga cgg cca gtc.

```

```

FEATURES
source
1..751
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="mRNA"
/cultivar="IR36"
/db_xref="taxon:39946"
/clone="OSIIIEa07L19"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSIIIEa"
/notes="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; Lesion Mimic SPL 11"

```

```

BASE COUNT      196 a 142 c 217 g 196 t
ORIGIN
Query Match      31.5%; Score 606; DB 14; Length 751;
Best Local Similarity 88.0%; Pred. No. 1.4e-70;
Matches 660; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

```

```

QY      331  TCAGTGTACACACAGCAGCAGTGGAGATCGCTGAAAACTTGTGTTGACAGGAACATGT 390
Db      2    TCAGTGTACACCCCTTACCAGTGGAGATCGCTGAGAACTTGTGTTGACAGGAACAT 61
QY      391  GGCCCCGTGCTATTCAAGAACTCACTAGCCTGAACCATATTGAAGAGGCTGCTGTCTTA 450

```

```

Db      62  GSCCCCGTGCTATCTCAGAACTCACCAGTCTGAATCATATTGAAGGGCTGCTGTTCTTA 121
QY      451  GTACCTGTAAATAGAAATTTATGTGTGGCGCTATCATGGAACCGTGTGTATCAGAG 510
Db      122  GTACCTGCAATAGAAATGGAATCTATGTGTAGCTTTTATCGTGGAAACCGTGGATTAGAG 181
QY      511  AAGTAGTGCAGTGGATGTGGAAGAAAGTGGTATTCCCGCTTCCGAGCTTACGGAGCAC 570
Db      182  AAGTGTAGATCTGGATGTCAAGAAAGTGGATCCCTGCTTCTGAGCTCAGGAGCATC 241
QY      571  TGTTCATCTTGCAGAGCAGTGTGACACACGCCATCTGTTTGGAGTGTGAGCTGGCTTG 630
Db      242  TATTCATGTTGCTGACAGTGTGACACACGCCATCTGTTTGGAGTGTGAGCTGGCTTG 301
QY      631  ACTCTTTGCTTCTCGTGAAGGACAAATCTTCTGCTCAGGTTAAACAAAGTTGTGAGAGTG 690
Db      302  ACTCTTTGCTTCTTGGAGAGGCGCAATCTTGTCTCAAGTTAAACAAAGTTGTGAGAGTG 361
QY      691  GACAGAACAGTGGAGGCTTGGGAAGAACATCGATAGGATGTTCAAGGATGCAATCAGTG 750
Db      362  GCGAANACAGTGGAGGCTTGGGAAGAACATCGATAGGATGTTCAAGGATGCAATCAGTG 421
QY      751  CTGGAAGCGTGTCCGACGAGACCAACATATCATCTGGTGTCTGTTCTGTGAGTTCAG 810
Db      422  CTGGAAGCGTGTCCGCTGCGAGCTAAACATATCATCATGAGTGTCTCTGTGAGTTCAG 481
QY      811  CGGCGGTTGAATCGGCCCTGATGAAGCTTCCGAAGCTTGAAGCACTGTGAGTAGATGC 870
Db      482  CTGCAAGTTGAATTTGGCTTGTATGAAGCTTCCAAAGTCGAATGCCCTATCTGCTAGATGC 541
QY      871  TTCTGATTGGTGGTAAATGGGAAGCTAGTATCAACATCTGCTTGCACCAAGGAT 930
Db      542  TGTGATTGGTGGTGAAGATGGGAAGTGGTGGTTAAACATTTGATTGCTCAAGGGAT 601
QY      931  GCAAGAGGTTGTTGTGTGAACCGCTCCGTGGAAAGGTTGATGATCTTCGTGAGGAGA 990
Db      602  GCAAGAGGTTGTTGTGTGAACCGTTCAGTGGAAAGGTTGATGATCCCGAAGAGA 661
QY      991  TGAAGATATAGAGATCGTGTACAGGCTCTCTCAGACATGTATCAAGCTGCTGTGAAG 1050
Db      662  TGAAGACATTCAGATTGTGTACAGGCTCTTACAGAGATGTATGAAGCCGCTGCCGAG 721
QY      1051  CTGATGCTGCTTCCACCAGCACCCCATCTG 1080
Db      722  CTGATGCTGCTTCCACCAGCACCCCATCTG 751

```

```

RESULT 14
AW520155/c
LOCUS      60061E06.xl 660 - Mixed stages of anther and pollen Zea mays cDNA,
DEFINITION mRNA sequence.
ACCESSION   AW520155
VERSION     AW520155.1 GI:7162533
KEYWORDS    EST.
SOURCE      Zea mays
ORGANISM    Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 608)
Walbot,V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu

```

```

FEATURES
  source
    Plate: 660061 row: E column: 06.
    Location/Qualifiers
      1..608
        /organism="Zea mays"
        /mol_type="mRNA"
        /cultivar="Ohio43"
        /db_xref="taxon:4577"
        /tissue_type="whole premeiotic anthers to pollen shed"
        /dev_stage="premeiotic anthers to pollen shed"
        /lab_host="XLOUR"
        /clone_lib="660 - Mixed stages of anther and pollen"
        /note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI;
        Site_2: XhoI; Anther and pollen cDNA library.
        Directionally sequenced with 5' end at the EcoRI site.
        Created by Amie Franklin."
      BASE COUNT      114 a 170 c 158 g 166 t
      ORIGIN
        Query Match      31.3%; Score 601.6; DB 9; Length 608;
        Best Local Similarity 99.3%; Pred. No. 5.9e-70;
        Matches 604; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1179 AATGTCAGCGCATGTGTCTGAAGTTGGCGCTGCACGAGTGTACAATGTCGACCACTTG 1238
Db |||||
608 AATGTCAGCGCATGTGTCTGAAGTTGGCGCTGCACGAGTGTACAATGTCGACCACTTG 549

QY 1239 AAAGAGGTGGTGAAGCCCAACAGGAGGACCGGCTCAGGAAGCAATCGAGCGCAGACA 1298
Db |||||
548 AAAGAGGTGGTGAAGCCCAACAGGAGGACCGGCTCAGGAAGCAATCGAGCGCAGACA 489

QY 1299 ATCATCACCAGAAACTGAGACGGTTTCGAGGATCGAGGGAATCCGCTCGAGACCCGTTCCG 1358
Db |||||
488 ATCATCACAAGAAGAACTGAGACGGTTTCGAGGATCGAGGGAATCCGCTCGAGACCCGTTCCG 429

QY 1359 ACCATCAAGAGCTGAGGTGTCACGGACAGGATCAGGCGCTCGAGCTCGAGAAGTGC 1418
Db |||||
428 ACCATCAAGAAGCTGAGGTGTCACGGACAGGATCAGGCGCTCGAGCTCGAGAAGTGC 369

QY 1419 CTGCAAGAAAGTGGTGAAGCGCCCTCACCAGAAAGATGAGGAGAGCCATCGAGAGCTG 1478
Db |||||
368 CTGCAAGAAAGTGGTGAAGCGCCCTCACCAGAAAGATGAGGAGAGCCATCGAGAGCTG 309

QY 1479 AGCACCGGCATCGTTAAACAAGCTCTCCATGCGCCGCTGCAGACCTGAGGTGCACGGC 1538
Db |||||
308 AGCACCGGCATCGTTAAACAAGCTCTCCATGCGCCGCTGCAGACCTGAGGTGCACGGC 249

QY 1539 AGCGACAGCGCACCCCTTGAAGAGAGCGCTCGAGAACATGCACGCCCTCAACCGGATGTC 1598
Db |||||
248 AGCGACAGCGCACCCCTTGAAGAGAGCGCTCGAGAACATGCACGCCCTCAACCGGATGTC 189

QY 1599 AGCCTCGACATGGAGAGCGCATCATCGAGCAGAAAGATCAAGGCCAAGGTGGAGAAGACA 1658
Db |||||
188 AGCCTCGACATGGAGAGCGCATCATCGAGCAGAAAGATCAAGGCCAAGGTGGAGAAGACA 129

QY 1659 CAAACTGAGGCGCAGAGAGCAATTTTCTACACCATTTATCTATATATATAGCGTCTCCA 1718
Db |||||
128 CAAACTGAGGCGCAGAGAGCAATTTTCTACACCATTTATCTATATATATATAGCGTCTCCA 69

QY 1719 ATCTCATTTCCATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1778
Db |||||
68 ATCTCATTTCCATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 9

QY 1779 TTAAGTGT 1786
Db |||||
8 TTAAGTGT 1

RESULT 15
CB685474
LOCUS
DEFINITION
  OSUNEF15004.f OSJNEF Oryza sativa (japonica cultivar-group) cDNA
  clone OSJNEF15004 5', mRNA sequence.
ACCESSION
  CB685474

VERSION
  CB685474.1 GI:29689199
EST.
  Oryza sativa (japonica cultivar-group)
  Oryza sativa (japonica cultivar-group)
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
  Ehrhartoideae; Oryzaceae; Oryza.
  1 (bases 1 to 733)
  Jantauriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
  Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
  Large-scale identification of ESTs involved in the interaction
  between rice and Magnaporthe grisea
  Unpublished
  Contact: Rod Wing
  Arizona Genomics Institute
  University of Arizona
  Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
  85721-0088, USA
  Tel: 520 626 3967
  Fax: 520 621 9288
  Email: http://genome.arizona.edu
  PCR Primers
  FORWARD: gta aaa cga cgg cca gtc
  BACKWARD: gga aac agc tat gac cat g
  Plate: 15 row: O column: 04
  Seq primer: gta aaa cga cgg cca gtc.
  Location/Qualifiers
    1..733
      /organism="Oryza sativa (japonica cultivar-group)"
      /mol_type="mRNA"
      /cultivar="Nipponbare"
      /db_xref="taxon:39947"
      /clone="OSJNEF15004"
      /tissue_type="Leaf"
      /dev_stage="3 week"
      /lab_host="DH10B"
      /clone_lib="OSJNEF"
      /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
      XhoI; Uninfected Control"
    BASE COUNT      193 a 140 c 209 g 191 t
    ORIGIN
      Query Match      30.4%; Score 584.2; DB 14; Length 733;
      Best Local Similarity 87.3%; Pred. No. 1e-67;
      Matches 640; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 409 AACTCAGTACGCTGAACCATATTGAAGAGCGTGTCTTTAGTACCTGTAAATAGAAATGG 468
Db |||||
1 AACTCAGTACGCTGAACCATATTGAAGAGCGTGTCTTTAGTACCTGTAAATAGAAATGG 60

QY 469 AAATTTATGCTGGCGCTATCATCGAACCGTGGTATCAGAGAGTAGTGGACTGGATGT 528
Db |||||
61 AAATTTATGCTGGTGTAGCTTTATCGTGAACCGTGGGATTTAGAGAGTGGTAGACTGGATGT 120

QY 529 CGAAGAAAAGTGTATTTCGCCGCTTCCGAGCTTAGGAGACCTGTTCATCTTGGGAACA 588
Db |||||
121 CAAAGAAAAGTGGAAATCCCTGCTTCTGAGCTCAGGAGCATCTATTATGTTGGGTGACA 180

QY 589 GTGATGCCACACGCCCATCTGTTTGGAGGTGTCAAGTGGCGCTTGAATCTTTTGGTTCGCGTG 648
Db |||||
181 GTGATGCCACACGCCCATCTGTTTGGAGGTATCTGCTGGGCTTGAATCTTTTGGTTCGCGTG 240

QY 649 AAGGACAAATCTTGTCTCAGGTTAAACAAGTTGTGAGAGGTGGAGACAGACAGTGGAGGCT 708
Db |||||
241 AAGGCAAAATCTTGTCTCAAGTTTAAACAAGTTGTGAGAGGTGGGCAAAAAGTGGAGGCT 300

QY 709 TGGGAAAAGAACATCGATAGGATGTTCAAGGATGCAATCAGCTGCGAAGCGTCTCCGCA 768
Db |||||
301 TGGGAAAAGAACATCGATAGGATGTTCAAGGATGCAATCAGCTGCGAAGCGTCTCCGCT 360

QY 769 GCGAGACCAACATATCATCTGTTGCTGTCTTCTGTGAGTTCAGCGCGCGGTGAACTGGCCCC 828
Db |||||
361 GCGAGACTAACATATCATCAGGTGCTGTCTCTGTGAGTTCAGCTGCGAGTGAATGGCCT 420
```

Qy	829	TGATGAAGCTTCCGAAGTCTGAAGCACTGTCAAGCTAGGATGCTCTCTGATTGGTGGCTGTA	888
Db	421	TGATGAAGCTTCCAAAGTCGGAATGCCTATCTGCTAGGATGCTGTGATTGGTGGTGGCA	480
Qy	889	AAATGGGAAGCTAGTGTCAAAACATCTGGTTGGCCAAAGGATGCAAGAGGTGTGTGG	948
Db	481	AGATGGGAAGTTGGTGGTTAAACATTTGATTGCCAAGGATGCAAGAAAGTTGTGG	540
Qy	949	TGAACCGCTCCGTGGAAAGGTGGATGCTATTCTGTGAGGATGAAAGATATAGAGTCG	1008
Db	541	TGAACCGTTCAGTGGAAAGGTGGATGCCATCCGGAAGAGATGAAGACATTGAGATTG	600
Qy	1009	TGTACAGGCTCTCTCAGACATGTATCAAGCTGTGCTGAAGCTGATGTCGTGTTTCA	1068
Db	601	TGTACAGGCTCTTACAGAGATGTATGAAGCGCTGCCGAAGCTGATGTGTGTTCAA	660
Qy	1069	GCACGGCATCTGAACCTTCTTTCGGCAAAAGAACACGCAAGGCACTCCCCCTGTCT	1128
Db	661	GCACGGCATCCGAACCCCATTTGTTTCAAAAGGAGCAGCAGAGCGCTTCCCGCTATT	720
Qy	1129	CTGATACTATGGG	1141
Db	721	CTGATGCTATGGG	733

Search completed: December 15, 2003, 16:07:25
Job time : 4199 secs

(DUFO) DU PONT DE NEMOURS & COE I.
XX
XX
Cahoon RE, Gutteridge S, Harvell LT, Rafalski JA, Tao Y, Weng Z;
PPI

XX WPI; 2001-159865/16.
DR P-PSDB; AAB31960.
XX New nucleic acids encoding an aminolevulinic acid biosynthetic enzyme
PT useful for producing large amounts of the encoded polypeptides used in
PT screening compounds for potential herbicidal activity -
XX
PS Claim 2; Page 47-48; 77pp; English.
XX The present sequence encodes a Glu-tRNA reductase, which is an
CC aminolevulinic acid biosynthetic enzyme. The enzyme converts
CC Glu-tRNA-Glu to glutamate 1-senialdehyde (GSA) with the concomitant
CC release of tRNA-Glu. GSA aminotransferase then converts GSA to
CC 5-aminolevulinic acid (ALA). ALA is used in the biosynthesis of
CC tetrapyrrolic pigments such as chlorophyll. Nucleic acids encoding
CC aminolevulinic acid biosynthetic enzymes are useful for isolating cDNAs
CC and genes encoding homologous proteins from the same or other plant
CC species, for creating transgenic plants in which the polypeptides are
CC present at higher or lower levels than normal or in cell types or
CC developmental stages in which they are not normally found, for
CC overexpression in bacterial or yeast hosts to efficiently produce
CC large amounts of the encoded polypeptides which could then be used for
CC screening different compounds for potential herbicidal activity, and
CC as hybridisation probes and amplification primers.
XX
SQ Sequence 1924 BP; 504 A; 465 C; 551 G; 404 T; 0 other;
Query Match 100.0%; Score 1924; DB 22; Length 1924;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1924; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCACGGTCCGATCAATTAAGAGGAGCTTGGAAAGTTGCCAAGGCTCCGATTTCGCTA 60
DB 1 CCACGGTCCGATCAATTAAGAGGAGCTTGGAAAGTTGCCAAGGCTCCGATTTCGCTA 60
QY 61 ATGCGACGATTAATGCGACACGACGCTGAGCCACCGCCGCGCAGCGCCGCCA 120
DB 61 ATGCGACGATTAATGCGACACGACGCTGAGCCACCGCCGCGCAGCGCCGCCA 120
QY 121 CGCCCAAGCGCGGGGTGCTGCTCGGCCCTCTGCGAGAGGTGGCGCGCGCGCTGG 180
DB 121 CGCCCAAGCGCGGGGTGCTGCTCGGCCCTCTGCGAGAGGTGGCGCGCGCGCTGG 180
QY 181 GCGCTCCGGGTGCTGCGGTGCGACGCGCGCGCGCTGAGGCGCGCGCGCGCTGG 240
DB 181 GCGCTCCGGGTGCTGCGGTGCGACGCGCGCGCGCTGAGGCGCGCGCGCGCTGG 240
QY 241 CAAAGGCTGCCAGCGCTCGCGCCCTCGAGCAGTTCAAGATATCCGCCACCGGTACATGA 300
DB 241 CAAAGGCTGCCAGCGCTCGCGCCCTCGAGCAGTTCAAGATATCCGCCACCGGTACATGA 300
QY 301 AGGAAAGAGTACCATAGCTGTGATAGGCTCAGTGATACACAGCACAGTGGAGATGC 360
DB 301 AGGAAAGAGTACCATAGCTGTGATAGGCTCAGTGATACACAGCACAGTGGAGATGC 360
QY 361 GTGAAAACTTGCTGTCAGAGGAACGTGCGCCCGCTGCTTCAAGAACTCACTAGCC 420
DB 361 GTGAAAACTTGCTGTCAGAGGAACGTGCGCCCGCTGCTTCAAGAACTCACTAGCC 420
QY 421 TGAACCAATTGAAGAGCTGCTGTTTCTTAGTACCTGTAATAGAAATTTATGTGG 480
DB 421 TGAACCAATTGAAGAGCTGCTGTTTCTTAGTACCTGTAATAGAAATTTATGTGG 480
QY 481 TGGCGCTATCATGACCGTGGTATCAGAGAAAGTAGTGGAATGATGTCGAAGAAAGTG 540
DB 481 TGGCGCTATCATGACCGTGGTATCAGAGAAAGTAGTGGAATGATGTCGAAGAAAGTG 540
QY 541 GTATTCCCGCTTCCGAGCTTAGGAGCACCTGTTTCATCTTCGAGCAGTATGCCAC 600
DB 541 GTATTCCCGCTTCCGAGCTTAGGAGCACCTGTTTCATCTTCGAGCAGTATGCCAC 600
QY 601 GCCATCTGTTTGAAGGTGTGAGCTGGCCCTTGACTCTTTGGTTCTCGGTGAAGCAAAATCC 660

DB 601 GCCATCTGTTTGAAGGTGTGAGCTGGCCCTTGACTCTTTTGGTTCTCGGTGAAGCAAAATCC 660
QY 661 TTGCTCAGGTTAAACAAGTTGCTGAGGAGTGCACAGAACAGTGGAGGCTTGGGAAAGAAC 720
DB 661 TTGCTCAGGTTAAACAAGTTGCTGAGGAGTGCACAGAACAGTGGAGGCTTGGGAAAGAAC 720
QY 721 TCGATAGGATGTTCAAGGATGCAATCACTGCTGGAAGCGGTGTCGCGAGCGAGCAACA 780
DB 721 TCGATAGGATGTTCAAGGATGCAATCACTGCTGGAAGCGGTGTCGCGAGCGAGCAACA 780
QY 781 TATCATCTGGTCTGTTCTGTCAGTTCAGCGGGGTGAACTGGCCCTGATGAAGCTTC 840
DB 781 TATCATCTGGTCTGTTCTGTCAGTTCAGCGGGGTGAACTGGCCCTGATGAAGCTTC 840
QY 841 CGAAGTCTGAAGCACTGTCACTAGGATGCTTCTGATTTGGTGGTAAATGGGAAAGC 900
DB 841 CGAAGTCTGAAGCACTGTCACTAGGATGCTTCTGATTTGGTGGTAAATGGGAAAGC 900
QY 901 TAGTGATCAACATCTGTTGCCAAAGGATGCAAGAGGTGTTGTTGTTGTAACCGCTCG 960
DB 901 TAGTGATCAACATCTGTTGCCAAAGGATGCAAGAGGTGTTGTTGTTGTAACCGCTCG 960
QY 961 TGGAAAGGCTGATGCTATTCTGAGGAGATGAAAGATATAGAGATCGGTACAGCCCTC 1020
DB 961 TGGAAAGGCTGATGCTATTCTGAGGAGATGAAAGATATAGAGATCGGTACAGCCCTC 1020
QY 1021 TCTCAGACATGTATCAAGCTCTCTGTAAGCTGATGCTGTTTCAACAGCAGCCGATCTG 1080
DB 1021 TCTCAGACATGTATCAAGCTCTCTGTAAGCTGATGCTGTTTCAACAGCAGCCGATCTG 1080
QY 1081 AAATCTTCAATTTGCGAAAGAACACGCGAGAGGCACTCCCGCTGCTCTGATATATGG 1140
DB 1081 AAATCTTCAATTTGCGAAAGAACACGCGAGAGGCACTCCCGCTGCTCTGATATATGG 1140
QY 1141 GAGGTGTTGCGCTGTTGTCGACATATCTGTCGCCAGGAATGTGAGCGCATGTGTTCTG 1200
DB 1141 GAGGTGTTGCGCTGTTGTCGACATATCTGTCGCCAGGAATGTGAGCGCATGTGTTCTG 1200
QY 1201 AAGTTGCGCTGCAAGTGTACAATGTGCAAGCTTTGAAAGAGGTGTTGGAAGCAACA 1260
DB 1201 AAGTTGCGCTGCAAGTGTACAATGTGCAAGCTTTGAAAGAGGTGTTGGAAGCAACA 1260
QY 1261 AGGAGGACCGGCTCAGGAAGCAATGAGGCGCAGACATCATCAACCGAAGATCTGAGAC 1320
DB 1261 AGGAGGACCGGCTCAGGAAGCAATGAGGCGCAGACATCATCAACCGAAGATCTGAGAC 1320
QY 1321 GGTTCGAGGCAATGAGGAGCTCGCTGAGACCGTTCGAGCCATCAAGAGCTGAGGTGCT 1380
DB 1321 GGTTCGAGGCAATGAGGAGCTCGCTGAGACCGTTCGAGCCATCAAGAGCTGAGGTGCT 1380
QY 1381 ACCTGAGGATCAGGCGCTCGAGGCTCGAGAGTGTCTGCAAGAAAGTAGGTGAGGACG 1440
DB 1381 ACCTGAGGATCAGGCGCTCGAGGCTCGAGAGTGTCTGCAAGAAAGTAGGTGAGGACG 1440
QY 1441 CCCTCAACAGAGATGAGGAGCGCATCGAGGAGCTGAGCACCGGCTTCGTTAACAAGC 1500
DB 1441 CCCTCAACAGAGATGAGGAGCGCATCGAGGAGCTGAGCACCGGCTTCGTTAACAAGC 1500
QY 1501 TCCTCCATGCGCCCTGCGAGCCTGAGGTCGAGCGCAGCGAGCGAGCGCCCTTGAAGC 1560
DB 1501 TCCTCCATGCGCCCTGCGAGCCTGAGGTCGAGCGCAGCGAGCGAGCGCCCTTGAAGC 1560
QY 1561 AGACGCTCGAGAACATGACCGCTCAACCGGATGTTTTCAGCTCGACATGAGAGGCGGA 1620
DB 1561 AGACGCTCGAGAACATGACCGCTCAACCGGATGTTTTCAGCTCGACATGAGAGGCGGA 1620
QY 1621 TCATCGAGCAGAGATCAAGCGCTGAGGAGAGACAACTAGGAGCGCAGGAGCA 1680
DB 1621 TCATCGAGCAGAGATCAAGCGCTGAGGAGAGACAACTAGGAGCGCAGGAGCA 1680
QY 1681 TTTTCTTACCACTTATCTATATATAGCTCTCAATCTCAATCTCAATCTTATTTTATCC 1740

Db 1681 TTTTCTACCACTTATCTATATATATAGCGTCTCCAATCTCAATTCATTTTATCC 1740
 QY 1741 TTTTCTACCTAGTGGAGCCCTCCCTGCTCACTGTGATCGTTAACTGTGTCTGTAATTAGA 1800
 Db 1741 TTTTCTACCTAGTGGAGCCCTCCCTGCTCACTGTGATCGTTAACTGTGTCTGTAATTAGA 1800
 QY 1801 GCCATGGCAGCGTGTGTCAATAAACAAGCAATGTGTCCCAATTCCTCCACAGAAAGAACT 1860
 Db 1801 GCCATGGCAGCGTGTGTCAATAAACAAGCAATGTGTGTCCCAATTCCTCCACAGAAAGAACT 1860
 QY 1861 ATATTATATGCAATTTATTTGAGCAAAATAGTTTACTTAAAAAATTTAAAAAATTTAAAAA 1920
 Db 1861 ATATTATATGCAATTTATTTGAGCAAAATAGTTTACTTAAAAAATTTAAAAAATTTAAAAA 1920
 QY 1921 AAAG 1924
 Db 1921 AAAG 1924

RESULT 2

AAF54816

ID AAF54816 standard; cDNA; 1778 BP.

XX AC AAF54816;

XX DT 15-MAY-2001 (first entry)

XX DE Nucleotide sequence of a rice Glu-tRNA reductase enzyme.

XX KW Glu-tRNA reductase; aminolevulinic acid; 5-aminolevulinic acid; ALA;

XX OS tetrapyrrolic pigment; chlorophyll; transgenic plant; ss.

XX FT Oryza sativa.

XX FT Key Location/Qualifiers

XX FT CDS 16..1458

XX FT /tag= a

XX FT /product= "Glu-tRNA reductase"

XX PN WO200109304-A2.

XX PD 08-FEB-2001.

XX PF 28-JUL-2000; 2000WO-US21008.

XX PR 30-JUL-1999; 99US-0146600.

XX PA (DUPO) DU PONT DE NEMOURS & CO E I.

XX PI Cahoon RE, Gutteridge S, Harvell LT, Rafalski JA, Tao Y, Weng Z;

XX DR WPI; 2001-159865/16.

XX DR P-PSDB; AAB31962.

XX PT New nucleic acids encoding an aminolevulinic acid biosynthetic enzyme

XX PT useful for producing large amounts of the encoded polypeptides used in

XX PT screening compounds for potential herbicidal activity

XX PS Claim 2; Page 52; 77pp; English.

XX CC The present sequence encodes a Glu-tRNA reductase, which is an

XX CC aminolevulinic acid biosynthetic enzyme. The enzyme converts

XX CC Glu-tRNA-Glu to glutamate 1-semialdehyde (GSA) with the concomitant

XX CC release of tRNA-Glu. GSA aminotransferase then converts GSA to

XX CC 5-aminolevulinic acid (ALA). ALA is used in the biosynthesis of

XX CC tetrapyrrolic pigments such as chlorophyll. Nucleic acids encoding

XX CC aminolevulinic acid biosynthetic enzymes are useful for isolating cDNAs

XX CC and genes encoding homologous proteins from the same or other plant

XX CC species, for creating transgenic plants in which the polypeptides are

XX CC present at higher or lower levels than normal or in cell types or

XX CC developmental stages in which they are not normally found, for

XX CC overexpression in bacterial or yeast hosts to efficiently produce

XX CC large amounts of the encoded polypeptides which could then be used for

CC screening different compounds for potential herbicidal activity, and
 CC as hybridisation probes and amplification primers.

XX Sequence 1778 BP; 496 A; 386 C; 479 G; 417 T; 0 other;
 SQ

Query Match 62.2%; Score 1196.4; DB 22; Length 1778;
 Beat Local Similarity 87.0%; Pred. No. 3e-249;
 Matches 1314; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 210 GCCGCGTGGAGGCCAGCGCAGCGCCTGGCAAGGCTGCCAGCGTCCGCCCTCGAG 269
 Db 1 GCACGAGTGGTACCCAGCGCAGCGGTGGCAAGGCGCGCAGCGTCCGCCCTCGAG 60
 QY 270 CAGTTCAAGATATCGCCGACCGGTACATGAAGGAAGAGTACCATAGCTGTGATGCC 329
 Db 61 CAGTTCAAGATATCGCCGACCGGTACATGAAGGAAGAGTACCATAGCTGTGATGCC 120
 QY 330 CTCAGTGTACACACAGCACCAGTGGAGATGCTGAAATCTTGCTGTTCGAGCAACTG 389
 Db 121 CTCAGTGTACACATGACACAGTGGAGATGCTGAAATCTTGCTGTTCGAGCAACTA 180
 QY 390 TGGCCCGTGTCTATTCAAGAACTCACTAGCCCTGAACCATATTGAAGAGGCTGCTCTT 449
 Db 181 TGGCCCGTGTCTATTCAAGAACTCACTAGCTGTGAATCATATTGAAGAGGCTGCTCTT 240
 QY 450 AGTACCTGTATAGAAATGAAATTTATGTGTGGCGCTATCATGGAACCGTGTATCAGA 509
 Db 241 AGTACCTGTAAATAGAAATTTATGTGTGGCGCTATCATGGAACCGTGTATCAGA 300
 QY 510 GAAGTAGTGGACTGGATGTCGAAGAAAGTGGTATTCGCCGCTTCGAGCTTAGGAGCAC 569
 Db 301 GAAGTAGTGGACTGGATGTCGAAGAAAGTGGTATTCGCCGCTTCGAGCTTAGGAGCAT 360
 QY 570 CTGTTCACTTTGCAAGCAGTGTATGCCACGCCATCTGTTTGAAGGCTTCAGCTGCCCTT 629
 Db 361 CTATTCATGTTGCGTGACAGTGTGCCACAGCCATCTGTTTGAAGGCTTCAGCTGCCCTT 420
 QY 630 GACTCTTTGGTCTCGGTGAAGGCAAAATCTTGTCTCAGGTAAACAAGTTGTGAGGAGT 689
 Db 421 GACTCTTTGGTCTCGGTGAAGGCAAAATCTTGTCTCAGGTAAACAAGTTGTGAGGAGT 480
 QY 690 GGACAGAACAGTGGAGGCTTGGGAAAGAAATCCATAGGATGTTTCAAGAGTGCATCACT 749
 Db 481 GGGCAAAACAGTGGAGGCTTGGGAAAGAAATCCATAGGATGTTTCAAGAGTGCATCACT 540
 QY 750 GCTGGAAGCGTGTCCGAGCGAGACCAACATATCATCTGCTGCTGTTTCTGTCAAGTTCA 809
 Db 541 GCTGGAAGCGTGTCCGAGCGAGACTAAACATATCATCAGGTGCTGTCTGTCAAGTTCA 600
 QY 810 CGCGCGGTGAACTGGCCCTGATGAAGCTTCGGAAGTCTGAAGCACTGTCAAGTATGAG 869
 Db 601 GCTGCAAGTTGAAATTTGGCTTTGATGAAGCTTCAAAAGTCCGAATGCTATCTGCTAGGATG 660
 QY 870 CTTCTGATTTGCTGCTGTTAAATGGGAAAGCTAGTGATCAACATCTGTTTGCACAAAGGA 929
 Db 661 CTGTTGATTTGCTGCTGCAAGATGGGAAAGTGTGTTGTTAAACATTTGATTTGCCAAGGA 720
 QY 930 TGCAGGAAGGTTGTTGTGTGAACCGCTCCGTCGGAAGGGTGGATGCTATTCTGTCAGGAG 989
 Db 721 TGCAGGAAGGTTGTTGTGTGAACCGCTTCAAGTGGAAAGGGTGGATGCTATCCGGAAGAG 780
 QY 990 ATGAAACATATAGAGATCGTGTACAGGCTCTCTCAGACATGATATCAAGCTGTGCTGAA 1049
 Db 781 ATGAAACATATGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
 QY 1050 GCTGATGCTGTTTCAACAGCAGCGCATCTGAAACTTCAATTTGTCGCAAGAAAGACGCA 1109
 Db 841 GCTGATGCTGTTTCAACAGCAGCGCATCTGAAACTTCAATTTGTCGCAAGAAAGACGCA 900
 QY 1110 GAGCACTCCCGCTGTCTCTGATATATGAGGAGGTTGTCGCTGTTTGTGCAATATCT 1169
 Db 901 GAGCGCTTCCCGCTATTCTGATGCTATGAGGTTGTTGCTGCTTCTGCTGCTGCTGCTGCT 960


```
Db 606 TGGTTGTCAACATTTGATTGCCAAGGATGCAAGAGGTTCTTGTGGTGAACCGTTCTG 665
QY 961 TCGAAGGCGTGCATCTATTCTGTGAGGAGATGAAGATATAGAGATCGTGTACAGCCCTC 1020
Db 666 TCGAAGGCGTGCATCTATTCTCCCAAGAGATGAAGATATGAGATTGTGTACAGCCCTC 725
QY 1021 TCTCAGACATGATCAAGCTGCTGTCTGAAGCTGATGTCGTGTTTCAACAGCAGCGATCTG 1080
Db 726 TTACAGAGATGATGAAGCGCTGCTGAAGCTGATGTCGTGTTTCAACAGCAGCGATCTG 785
QY 1081 AAATTCATGTTTGGCAAAAGAACACGACAGAGGACTCCCCCTGCTCTGTGATCTATGG 1140
Db 786 AATCTCTATTATTACGAAGAGCATGACAGAGCGCTTCTCTCTATTCTCTGTGTGG 845
QY 1141 GAGGTGTTCCGCTGTTTGTGACATATCTGTCAGGATCTGTCAGGATCTGTCAGGATCTGTCG 1200
Db 846 GTGGTGTTCGGCTGTTTGTGACATATCTGTCAGGATCTGTCAGGATCTGTCAGGATCTGTCG 905
QY 1201 AAGTTGGCGCTGCACGAGTGTAACATGTCGACGACTTGAAGAGGTGGTGAAGCAACA 1260
Db 906 AGGTGAGCATGCACGGGTATACATGTCGACGACTTGAAGAGGTGGTGAAGCAACA 965
QY 1261 AGGAGACCGGCTCAGGAAAGCAATGAGGCGACGACAAATCATCCGGAAGAACTGAGAC 1320
Db 966 AGGAAGACCGTGTGAGGAAAGCAATGAGGCGCCCAACAAATCATCCGGAAGAACTGAGAC 1025
QY 1321 GGTTCGAGGATGAGGAGCTCGCTGGAGACCGTTCGACCACTCAAGAGCTGAGGATCTG 1380
Db 1026 GGTTCGAGGATGAGGAGCTCACTGGAGACCGTTCGACCACTCAAGAGCTGAGGATCTG 1085
QY 1381 ACGCGACAGGATCAGGCGCTCGGAGCTCGAAGTGCCTGCAGAAAGTAGGTGAGGAGC 1440
Db 1086 ACGCGACAGGATCAGGCGCTCGGAGCTCGAAGTGCCTGCAGAAAGTAGGTGAGGAGC 1145
QY 1441 CCCTCAGCAAGATGAGGAGCATCGAGGAGCTGAGCAACCGGCTATCGTTAAACAAGC 1500
Db 1146 ATCTCAACAAGAGATGAAGGTCCATCGAGGAGCTGAGCAACCGGCTATCGTTAAACAAGC 1205
QY 1501 TCCTCATGCGCGCTGACGACCTGAGGTGCGAGCGGACGACGACGCGCCCTTGGC 1560
Db 1206 TCCTTCAGCGCCACTGACGACCTGAGATGCGAGCGGACGACGACGCGCCCTTGGC 1265
QY 1561 AGACGCTCGAGAACATGACGCGCTCAACCGGATGTTGAGCCTCGACATGAGAGAGGGA 1620
Db 1266 AAACGCTTGAGAACATGACGCGCTCAACAGAAATGTTCAACCTCGACGAGAGAGCGG 1325
QY 1621 TCATCGAGCAGAGATCAAGGCAAGGTGGAGAGACACAACTGAGGCCAGGAGCAA 1680
Db 1326 TCCTTGAGCAGAGATCAAGGCAAGGTAGAGAGACCAAAAGCTGAGACCAAGAGACAC 1385
QY 1681 TTTTCTTACCACTTATCTATATATATAGCGTCTCCCAATCTCATTCATTTT 1733
Db 1386 TTGCCCGTC--TGTATATCTACTTATATCTGCTCCAGAAATGTCGTACATCT 1436
```

RESULT 4

ID AAF54818 standard; cDNA; 2055 BP.

AC AAF54818;

DT 15-MAY-2001 (first entry)

DE Nucleotide sequence of a soybean Glu-tRNA reductase enzyme.

KW Glu-tRNA reductase; aminolevulinic acid; 5-aminolevulinic acid; ALA; tetrapyrrolic pigment; chlorophyll; transgenic plant; ss.

OS Glycine max.

FH Key Location/Qualifiers

FT CDS

111..1718

/*tag= a

/product= "Glu-tRNA reductase"

WO200109304-A2.

08-FEB-2001.

28-JUL-2000; 2000WO-US21008.

30-JUL-1999; 99US-0146600.

(DUPO) DU PONT DE NEMOURS & CO E I.

Cahoon RE, Gutteridge S, Harvell LT, Rafalski JA, Tao Y, Weng Z;

WPI; 2001-159865/16.

P-PSDB; AAB31964.

New nucleic acids encoding an aminolevulinic acid biosynthetic enzyme useful for producing large amounts of the encoded polypeptides used in screening compounds for potential herbicidal activity -

Claim 2; Page 55-57; 77pp; English.

The present sequence encodes a Glu-tRNA reductase, which is an aminolevulinic acid biosynthetic enzyme. The enzyme converts Glu-tRNA-Glu to glutamate 1-semialdehyde (GSA) with the concomitant release of tRNA-Glu. GSA aminotransferase then converts GSA to 5-aminolevulinic acid (ALA). ALA is used in the biosynthesis of tetrapyrrolic pigments such as chlorophyll. Nucleic acids encoding aminolevulinic acid biosynthetic enzymes are useful for isolating cDNAs and genes encoding homologous proteins from the same or other plant species, for creating transgenic plants in which the polypeptides are present at higher or lower levels than normal or in cell types or developmental stages in which they are not normally found, for overexpression in bacterial or yeast hosts to efficiently produce large amounts of the encoded polypeptides which could then be used for screening different compounds for potential herbicidal activity, and as hybridisation probes and amplification primers.

Sequence 2055 BP; 596 A; 419 C; 465 G; 575 T; 0 other;

Query Match

Best Local Similarity 36.9%; Score 710.4; DB 22; Length 2055;

Matches 1019; Conservative 0; Mismatches 471; Indels 6; Gaps 2;

QY 189 GGGGTGTTGCGTGGCGACGCCCGCGGTGGAGGCCAGCGCGTGGCAAGGCT 248

Db 252 GGGGTATTGCTGCGACGCTCAGCCCTCTGATGATCATCTGTTGCTCCAAATATGCC 311

QY 249 GCCAGCTGCGCGCCCTCGACAGTTCAGATAT---CGCGGACCGGTACATGAAGGAA 305

Db 312 ACCGCTCTCTCCGCTCTTGAGCAGCTCAAGACTTCTGCAGCTGATAGATATACAAAGGAA 371

QY 306 AGGAGTACCATAGCTGTGATAGGCTCAGTGTACACAGCAGCAGTGGAGTGGTGA 365

Db 372 AGAGCAGCATATTCGCCATTTGGGCTCAGTGTGACACTGACCTGTGGAATGCGTGA 431

QY 366 AAACCTGCTGTGACAGAGAACTGTGGCCCGTGTCTATTCAAGAACTCACTAGCTGAAC 425

Db 432 AAACCTTGCCATTCCAGAGCAGAAATGGCTAGAGCTATTGCGAGCTGTGTAGTCTGAAT 491

QY 426 CATATTGAGAGGCTGTGTTCTTAGTACCTGTATAGAAATGGAATTTATGTGTGGGG 485

Db 492 CATATTGAAGAGCAGCTGTTCTGAGTACCTGCAATCGAATGGAGATATATGTTCTTGC 551

QY 486 CTATCATGGAACCGTGGTATCAGAGAACTGTGACTGTGCTGATGTCGAAGAAAGTGTATT 545

Db 552 CTGTCCCAACATCTGTGTCTCAAGAACTCATGGAATGGATGTCAAAAACAGTCTCTGC 611

QY 546 CCCGCTTCGAGCTTAGGAGCAGCTGTTCATCTTTGCGAAGCAGTGTGTCACACAGCCAT 605

Db 612 CCTGTTTCAGAGCTTAGCCAGCAGCGGTTTTTACTTTTACACAAATGATGCCACAGCAT 671

606 CTGTTTGAGGTGTCAGCTGGCTTGTACTCTTTGGTTCTCGGTGAAGGACAAATCCTTGT 665
672 CTTTTTGAAGTATCAGCAGGCTTGTACTCTCTTTGTTGGGGAAAGGTCAAAATCCTTTCT 731
666 CAGGTTAAACAAAGTTGTGAGGAGTGCACAGAACAGTGGAGGCTTGGGAAAGAACATCGAT 725
732 CAGGTTAAGCAAGTTGTTAAAGTTGACCAAGAGTTAACGGCTTTGGGAGAAATATCAGT 791
726 AGGATGTTCAAGATGCAATCACTGCTGGAAGCGTGTCCGAGCGAGACCAACATATCA 785
792 GGGCTATTCAAGCATGCAATTAATGTCGGGAAAGGGTTAGAACTGAGACTAATATGCT 851
786 TCTGGTGTCTGTTCTGTCAGTTTCAGCGCGGTTTGAACCTGGCTGATGAAGCTTCCGAG 845
852 TCTGGGCGAGTTTCTGTGAGCTCAGCTGCGTGTGAGTTGGCTTATATGAAGTTACTGAA 911
846 TCTGAAGCACTGTCACTAGGATGCTTCTGATTGTTGCTGGTAAATGGAAAGCTAGTG 905
912 GCCTCACACGATAATGCCAGGATGTTGGTTATTGTTGCTGGCAAGATGGAAAGCTTGTG 971
906 ATCAACATCTGGTTGCCAAAGATGCAAGAGGTTGTTGTTGGTGAACCGCTCCGTGGAA 965
972 ATCAACATCTGGTTGCCAAAGGTTGCAAAAAGATGGTGGTTGTCAATAGAACTGAGGAG 1031
966 AGGGTGGATGCTATTCTGTGAGGAGATGAAAGATATAGAGATCGTGTACAGGCTCTCTCA 1025
1032 AGAGTTGCTGCAATAGTCAAGACTGAAGGATATGAGATTATCTACAAACCCCTTCA 1091
1026 GACATGATCAAGCTGCTGCTGAAGCTGATGTCGTTTCCAGCAGCACCGCATCTGAAC 1085
1092 GAAATGCTCACCCTGTCTGGCAAGCAGATTAGTTTCCACCAGTACTGCATCAGAAAC 1151
1086 TCATTGTTCCGAAAGAACACGAGGACCTCCCGCTGCTCTGATACTATGGAGGT 1145
1152 CCATTATTCTGAAAGAAATGTCGAGGACCTTCTCTGCAAGTCAAGAGTTGGAGC 1211
1146 GTTCGGCTGTTTGTGACATATCTGTCCCGAGGATGTGAGCATGTGTCTGAAGTT 1205
1212 CGTCGCTTTTTCATTGATATCTGTGTTCCCGGATGTGGTTTCAATGTCTCAGACCTT 1271
1206 GGGCTGCAAGGATGCAATGTCAGCATCTGAAAGAGGTGGTGAAGCCCAACAGGAG 1265
1272 GAGTCTGTGCGAGTTTACATGTTGACGACCTTAAAGAGGTTGTGGCTGCCAATAAGAG 1331
1266 GACCGCTCAGAAAGCAATGAGGCGCAGACATCATCAGCAAGTCTGAGCGGTT 1325
1332 GATCGCTAAGAAAGCAATGGAAGCAGGCAATCATTCGTAAGAAATCTAAGCAATT 1391
1326 GAGGCAATGAGGAGCTCGCTGGAGACGTTCCGACCATCAAGAAGCTGAGGTGCTACGG 1385
1392 GAAGCTTGAGGAGCTCACTGGAAACTGTTCTTACTATTAAAGAAATGAGGCTTATGCT 1451
1386 GACAGATCAGGCTCGAGCTCGAAGTGTGCTGAGAAAGTGGTGAAGCCCTTC 1445
1452 GAAAGAAATCAGGCTTGTCTGAGTGTGAGAAAGTGTGTTAGTAAAGTGGTGATCA---TATA 1508
1446 ACCAAGAATGAGGAGCCATCGAGGAGCTGAGCACCGCATCGTTAAGCTCTTC 1505
1509 CCAAGAAACCGGAGACCTGTGATGACCTTGTGCGGGTATAGTGAATAGTTGCTT 1568
1506 CATGCGCCCTGCAGCCTGAGTGTGAGTGTGAGCGCAGCAGCGCACCCCTTGACGAGCG 1565
1569 CATGTGCCAATGCAATTTAAGTGTGATGGAGACGACAGCGGACCTTCTAGTGAGCA 1628
1566 CTCGAGAAATGCAAGCCCTCAACCGGATGTTGAGCTGCAATGAGGAGCGGATCATC 1625
1629 CTGGAGAAATGAATGCTTTGAATAGGATGTTCAACCTTGGACAGAAATATCTGTTTG 1688
1626 GAGCAGAAATCAAGCCCAAGTGGAGAGACACAAAACCTGAGGCCAGCAAT 1681
1689 GAGGAGAGATTCGAGCAAGGTCGAAACAAACCAAGAAATGAATCTAACACCAAT 1744

RESULT 5
ABZ13396
ID ABZ13396 standard; DNA; 1632 BP.
XX
AC ABZ13396;
XX
DT 21-JAN-2003 (first entry)
XX
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1201.
XX
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200216655-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US26685.
XX
PR 24-AUG-2000; 2000US-227866P.
PR 26-JAN-2001; 2001US-26467P.
PR 22-JUN-2001; 2001US-30011P.
XX
PA (SCEI) SCRIPPS RES INST.
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Harper JF, Kreps J, Wang X, Zhu T;
XX
WPI; 2002-304127/34.
XX
PT Identifying a stress condition to which a plant cell has been exposed
PT and producing plants with increased tolerance to these abiotic stresses
PT
XX
PS Claim 144; SEQ ID NO 1201; 577pp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising:
CC (a) contacting nucleic acid representative of expressed polynucleotides
CC in the plant cell with an array or probes representative of the plant
CC cell genome; and
CC (b) detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention.
CC Note: The sequence data for this patent is not represented in the printed
CC specification but is based on sequence information supplied to Derwent by
CC the European Patent Office.
XX
SQ Sequence 1632 BP; 463 A; 289 C; 428 G; 452 T; 0 other;
Query Match 35.3%; Score 679.6; DB 24; Length 1632;
Best Local Similarity 68.6%; Pred. No. 2.3e-137;
Matches 985; Conservative 0; Mismatches 439; Indels 12; Gaps 3;
QY 235 CCGTGGCAAGAGCTGCCAGCGCTCGCGCCCTCGAGCAGTTCAGAGATAT---CCGCGGACC 291
Db 206 CTGCGTCTAATGCAGCTAGCATCTGCTCTTGAACAACCTCAAGAACTCTGCAGCTGATC 265
QY 292 GGTACATGAAGAAAGGAGTACCATAGCTGTGATAGGCTCAGTGTACACAGACAGCAG 351
Db 266 GATATACAAAGGAAAGACAGTATTTGTTGATGGACTTAGTATTACACAGCTCCTG 325
QY 352 TGGAGATCGGTGAAGAACTTCTGTCAGAGGAACTTGGCCCGCTGCTATTCAAGAAC 411
Db 326 TTGAGATCGGTGAAGAACTTCTGATACCTGAAGCTTGAATGGCCAAAGAGCTATTGCTGAAT 385
QY 412 TCACCTAGCTGAACCATATTGAAGAGGCTGCTGTTCTTAGTACCTGTATAGAAATGGA 471
Db 386 TGTGTGTTGAATCATATTGAAGAGCTGCTGCTGTAGTACTTGTAAACGTATGGAGA 445

QY 472 TTTATGTGCTGGCGCTATCATGGAACCGTGGTATCAGAGAGTAGTGGACTGGATGTGCA 531
 Db |||||
 QY 446 TTTATGTTTATCTTATCTCAGCATCGTGGAGTTAAAGAGTTACTGAATGGATGTCAA 505
 Db |||||
 QY 532 AGAAAGTGGTATTCCTCGCTCCGAGCTTAGGGAGCACCTGTTTCATCTTGGCAAGCAGTG 591
 Db |||||
 QY 506 AGACGAGTGGATCCAGTTTCAGAGATTGTCAACACCGTTTCTGTTGTACACAGG 565
 Db |||||
 QY 592 ATGCCACAGCCATCTGTTTGGTGTAGCTGGCTTGAAGTGTGCTTGGTTCGGTGAAG 651
 Db |||||
 QY 566 ATGCTACACAGCATATATTTGAAGTCTCAGCTGGTCTGACTCTCTGTTCTCCTAGGAG 625
 Db |||||
 QY 652 GACAAATCTGCTCAGGTTAAACAAAGTTGTGAGAGTGCAGACAGTGGAGGCTTGG 711
 Db |||||
 QY 626 GTCAGATCTTGCACAGGTGAACAAAGTTGTGAAGTTGGTCAAGAGTGAATGGCTTGG 685
 Db |||||
 QY 712 GAAGAATCATGATAGGATGTTCAAGGATGCAATCACTGCTGGAAAGCGTCCCGAGG 771
 Db |||||
 QY 686 GGAGGAATATCAGCGGCTGTTTAAACACGCGATAACTGTTGTAAGCGTGTAGAACAG 745
 Db |||||
 QY 772 AGACCAACATATCATCTGCTGCTGTTCTGTCAGTTCAGCGCGGTTGAATGCGCCCTGA 831
 Db |||||
 QY 746 AGACAAATATTTCTGCTGCGGTTTCGTTAGCTCAGCTGCTGTTGAATGCTCTGA 805
 Db |||||
 QY 832 TGAAGCTCCGAAAGTCTGAAGCACTGTCAGCTAGGATGCTTCTGATTTGCTGCTGTA 891
 Db |||||
 QY 806 TGAAGCTCCGAGTCTTCAACAGTATCTGCTAGGATGTTGTAATGCGCTGGAA 865
 Db |||||
 QY 892 TGGGAAGCTAGTGTCAACATCTGTTTGCACAAAGGATGCAAGAGGTTGTTGGTGA 951
 Db |||||
 QY 866 TGGGAAGCTTGTGATTAAGCATTTGATGGCTAAGGTTGCACAAAGTGGTGTAGTCA 925
 Db |||||
 QY 952 ACCGCTCGTGAAGGTTGATGCTATTCGTGAGAGATGAAGATATAGAGTCGTGT 1011
 Db |||||
 QY 926 ACAGAAGTGAAGAGAGGTTTCAGCTATCCGTGAGGAGATGCTGCTGTTGAGATTATAT 985
 Db |||||
 QY 1012 ACAGGCTCTCTCAGACATGATCAAGCTGCTGCTGAAGCTGCTGCTGCTGCTGCTG 1071
 Db |||||
 QY 986 ATCGACCTCTTGATGAGATGCTAGCTTGTGCTTCAAGAGCGATGTTGTTTACCAGCA 1045
 Db |||||
 QY 1072 CCGCATCTGAACCTTCACTTTCGCAAAAGAACACGACGAGGCACTCCCGCTGCTCTG 1131
 Db |||||
 QY 1046 CAGCTCTGAGAGCGCATTTCTTAAAGGAGCATGTAGAGATCTCCCTCAAGCTTCTC 1105
 Db |||||
 QY 1132 ATACTATGGAGGTTTGCCTGTTTGTGCAATATCTGTCGCCAGGAATGTACGCGAT 1191
 Db |||||
 QY 1106 CAGAGGTTGGAGGACTGAGGCAATTTTGTGACATCTCTGTTCCGAGAAATGTTGATCTT 1165
 Db |||||
 QY 1192 GTGTGCTCTGAAGTTGGCGCTGCACGAGTGTCAATGTCGACGACTTGAAGAGGTTGG 1251
 Db |||||
 QY 1166 GTGTTGGTGAAGTAGAAGACAGCAGGTTTACATGTGAGCGATCTCAAGAGTTGTTG 1225
 Db |||||
 QY 1252 AAGCCAAAGAGAGACCGGCTCAGAAAGCAATGGAGGCGCAGCAATCATCATCCGAG 1311
 Db |||||
 QY 1226 CTGCCAAACAAAGACAGGATGAGAAAGCAATGGAAGCTCAAAACCATATCATCAGAG 1285
 Db |||||
 QY 1312 AACTGAGACGTTGAGGATGAGGAGTCTGCTGGACACGTTCCGACCATCAAGAGC 1371
 Db |||||
 QY 1286 AATCCACAGTTTGAACATGAGGAGGATTCATTGGAGACGTTCCACGATCAAGAGT 1345
 Db |||||
 QY 1372 TGAGTCTGACCGACAGGATCAGGGCTCGAGCTCGAAGTGGCTCGAGAGTGAAGTAG 1431
 Db |||||
 QY 1346 TAAGAGCTTATGACAGAGATCAGATGGGAGAGCTCGAAGTGCATGTCCAAATGG 1405
 Db |||||
 QY 1432 GTGAGGAGCCCTCACCAGAAAGATGAGAGAGCCATCGAGGAGCTGAGCAACCGGCATCG 1491
 Db |||||
 QY 1406 GAGATGA---TATCAACAAGAAACACAGAGCGGTTGATGACTTAAGCCGAGGTATCG 1462
 Db |||||
 QY 1492 TTAACAAGCTCTCCATGGCCGCTGAGACCTGAGTGGAGCGGAGCGAGCGACCGCA 1551
 Db |||||
 QY 1463 TGAACAGATTTTACATGTTCCAAATGCGAGCATTTTGATGTGTGACGGAAGTGACAGTAGAA 1522
 Db |||||

QY 1552 CCCTTGACGAGAGCGCTCGAGACATGCGCCCTCAACCGGATGTTTCAGCTCGCATGG 1611
 Db |||||
 QY 1523 CGCTGAGGAGACCTTGGCATATATGCAATGCTTTGACACAGAAATGTACCGTC-----TTG 1576
 Db |||||
 QY 1612 AGAAGCGCATCATCGAGCAGAGAGATCAAGGCCAAGTGGAGAGACACAAAACCTGA 1667
 Db |||||
 QY 1577 AGAAGAGACATTTTGGAGGAAAGCTTAAGGCAATGGCGAACACACAGAGAGTAA 1632
 Db |||||

RESULT 6

AAF54820
 ID AAF54820 standard; cDNA; 1983 BP.

XX AAF54820;

XX AC AAF54820;

XX DT 15-MAY-2001 (first entry)

XX Nucleotide sequence of a soybean Glu-tRNA reductase enzyme.

XX Glu-tRNA reductase; aminolevulinic acid; 5-aminolevulinic acid; ALA;

XX tetrapyrrolic pigment; chlorophyll; transgenic plant; ss.

XX Glycine max.

XX Key Location/QualifierB

XX CDS 289..1693

XX FT /*tag= a

XX FT /product= "Glu-tRNA reductase"

XX PN WO200109304-A2.

XX XX 08-FEB-2001.

XX PD 28-JUL-2000; 2000WO-US21008.

XX PF 30-JUL-1999; 99US-0146600.

XX PR (DUPO) DU PONT DE NEMOURS & CO E I.

XX PA Cahoon RE, Gutteridge S, Harvell LT, Rafalski JA, Tao Y, Meng Z;

XX XX WPI, 2001-159865/16.

XX DR P-P8DB; AAB31966.

XX XX New nucleic acids encoding an aminolevulinic acid biosynthetic enzyme

XX PT useful for producing large amounts of the encoded polypeptides used in

XX PT screening compounds for potential herbicidal activity -

XX XX Claim 2; Page 62; 77pp; English.

XX The present sequence encodes a Glu-tRNA reductase, which is an

XX aminolevulinic acid biosynthetic enzyme. The enzyme converts

XX Glu-tRNA-Glu to glutamate 1-semialdehyde (GSA) with the concomitant

XX release of tRNA-Glu. GSA aminotransferase then converts GSA to

XX 5-aminolevulinic acid (ALA). ALA is used in the biosynthesis of

XX tetrapyrrolic pigments such as chlorophyll. Nucleic acids encoding

XX aminolevulinic acid biosynthetic enzymes are useful for isolating cDNAs

XX and genes encoding homologous proteins from the same or other plant

XX species, for creating transgenic plants in which the polypeptides are

XX present at higher or lower levels than normal or in cell types or

XX developmental stages in which they are not normally found, for

XX overexpression in bacterial or yeast hosts to efficiently produce

XX large amounts of the encoded polypeptides which could then be used for

XX screening different compounds for potential herbicidal activity, and

XX as hybridisation probes and amplification primers.

XX Query 1983 BP; 514 A; 413 C; 479 G; 577 T; 0 other;

Query Match

Best Local Similarity 30.6%; Score 589.2; DB 22; Length 1983;

Matches 902; Conservative 0; Mismatches 473; Indels 9; Gaps 2;

QY 282 TCCGCCACCGGTACATGAAGGAAAGAGTACCATTAGCTGTGATAGCCCTCAGTGTACAC 341


```

||||| 317 TCCGCGCAGATATACGAGGAAAGAGTTGCATATTTGCTAGGCGTGAACATTCAC 376
QY 342 ACAGCACCAGTGGAGATCGGTGAAATACTGCTGTTCAGAGGAACTGTGCCCCGTGCT 401
Db 377 ACTGCTCCCGTTGAGATCGTGAAGAGCTTGCATATTCAGAAATCCCATTTGGGCTCAGGCT 436
QY 402 ATTCAAGAACTCCTAGCTGAACCATATTTGAAGAGGCTGTCTTTAGTACCTGTAAT 461
Db 437 ATTAAGGACCTTTGGCTTTGAAACCATATCGAAGAGCGCGTTCTCAGCAGCTGTAAAC 496
QY 462 AGAATGGAATTTATGTGGCGCTATCATGGAACCGTGTATCAGAGAAAGTAGTGGAC 521
Db 497 CGCATGGAGATCTATGTTGGCTCTTTCCAGCACCGTGTGTTTAAAGAAAGTTACTGAT 556
QY 522 TGGATGTCGAAGAAAGTGTATTCGCCGTTCCGAGCTTAGGAGCACCTGTTCACTTCTG 581
Db 557 TGGATGTCGAAGAGTGAAGGAGATTTCAATACCTGAGCTTTGTGAGCACCAAGTTTGGCTG 616
QY 582 CGAAGCAGTGTGCGCACAGCCATCTGTTTGAAGTGTGAGCTGGCGCTTGACTCTTTGGTT 641
Db 617 TATAAGCGGATGTACGCGCATCTCTTTGAGTGGCGGAGGCTTGACTCACTTGT 676
QY 642 CTCGTGAAGGACAAATCTTGTCTCAGGTTAAACAAAGTTGTGAGGAGTGGACAGAACAGT 701
Db 677 CTGGGGAAGGTCAAAATCTTGTCTCAGGTGAAGCAGGTTGTGAAAGCTGGACAGGAGTG 736
QY 702 GGAGCTTTGGGAAGAAACATCGATAGGATGTTCAAGGATGCAATCACTGCTGGAAGCGT 761
Db 737 CCGTGTGTTGTAAGAAATTTAGTGGTTTGTTCAGCAGGCGATCTCGGTTGGGAAGCGG 796
QY 762 GTCCGACGAGACCAACATATCATCTGCTGCTGTTCTGTCAGTTTCAAGCGCGGTTGAA 821
Db 797 GTTAGAAGTGAAGTAAATCTGCTGAGTGGGTTCTGTCAGCTCGGCTGCTGTGGAG 856
QY 822 CTGGCCTGTAGTGAAGCTTCCGAGTCTGAAGCTGTGAGTAGGATGCTTCTGATGCT 881
Db 857 CTCGCACTGATGAAGCTTCCGAGTCTCTCTTGTCTGATTTCTGGAGTGTGCTGTTGGT 916
QY 882 GCTGTAAAGTGGGAAGCTAGTATCAAACTCTGTTGTCGAAGGATGCAAGAGTT 941
Db 917 GCAGGAGATGGGAAGCTTGTAAATGAAGATTTGGCTGCAAGGAGTGCAGAGAAATG 976
QY 942 GTTGTGTGAACCGCTCCGTTGGAAGGTTGATGCTATTCGTGAGGAGATCAAGAGATATA 1001
Db 977 GTTGTGTTAAGCAGACTGAAGAGAAAGTTATGCAATTCGAAGAGTTGAAGATGTT 1036
QY 1002 GAGATCGTGAAGCGCTCTCTCAGACATGTATCAAGCTGTGCTGAGAGTGTGCTG 1061
Db 1037 GAGATGTTATTTAGACCATTTTCAAGATATGCTGGCGTGTGCTGAGAGTGTGATC 1096
QY 1062 TTCACGACCGCATCTGAAGCTTCAATGTTTCGAAGAACACGAGAGGCTCCCT 1121
Db 1097 TTCACGACGACGCTTGAATCACCATTGTTTCTAAACAGAAATGTCAGATGCTCTCT 1156
QY 1122 CCTGTCTCTGATATGAGGAGTGTTCGCTGTTTGTGACATATCTGCTCCCGAGAAAT 1181
Db 1157 CTGCT-----TAATCATGGGAGAGGCGGCTTTTGTGATATCTATTCCTAGAGAT 1210
QY 1182 GTCAGCGCATGTGTCTGAAGTTGGCGCTGCAAGAGTGTACAATGTCAGCACTTGAAA 1241
Db 1211 GTGGAACCGGCTGTCTCAGATCTCGAGACTGCACTTGTGTACAATGTCAGATGATCTGAAG 1270
QY 1242 GAGGTGTGGAAGCAACAGAGGACCGGCTCAGGAAGCAATGAGGCGGACACATC 1301
Db 1271 GAAGTTGTGAGCTTAACAGAGAGACAGGCTTCAGAAAGCTGAGGAAGCCCGGGAAT 1330
QY 1302 ATCACCAGAACTGAGACCGTTTCAGGCAATGGAGGAGCTCGCTGGAGACCGTTCCGACC 1361
Db 1331 ATACTAGAGGAGTTGAATTAATTCGAAGCTTGGAAAGACTCTCTGGAACCTGTTCTTACT 1390
QY 1362 ATCAGAGCTGAGGTGTGACCGGACAGGATCAGGCGCTCGAGCTCGAGAAAGTGCCTG 1421

```

```

Db 1391 ATTAAGAGTTTAGAGCTTATGTTGAGAGGATAAGAGCCTCTGAGATGAGAGTGTGTTG 1450
QY 1422 CAGAAAGTAGTGAGGAGCGCCTCACCAAGAAAGTAGAGAGGAGCCATCGAGAGCTGAGC 1481
Db 1451 TCGAAGATGGGTCTGATG---TCTCAAGCAACAGAAAGATGCAATTTATGCCCTTAGT 1507
QY 1482 ACCGCGCATCGTTAAACAGCTCTCCATGCGCCGCTGCGAGCACCTGAGGTGGACGCGAGC 1541
Db 1508 ATGGGTATTTGTAATAGCTACTTCTATGCTGCCATGCGAGCACCTTAAGGTGTGATGGGAA 1567
QY 1542 GACAGCGCACCCCTTGAGAGAGCCTCGAGAACATGCAAGCCCTCAACCGGATGTTTCAGC 1601
Db 1568 AATGATAGTAGTCTCTGAGTGAGGTACTTGAAGATATGCGTGCCCTTAACAGAAATGTACGAT 1627
QY 1602 CTCACATGGAAGAGCGATCATCGAGCAGAAAGATCAAGGCCAAGGTGGAGAAAGACAA 1661
Db 1628 CTTGAGACAGAAATTTCTCTTGTATCGAAGAAAGATCAGAGTCAAGATGGAACGGGTTTCAG 1687
QY 1662 AACT 1665
Db 1688 AAGT 1691

```

RESULT 7

```

AAF54815
ID AAF54815 standard; cDNA; 510 BP.
XX AAF54815;
XX 15-MAY-2001 (first entry)
XX Nucleotide sequence of a rice Glu-tRNA reductase enzyme.
XX
XX Glu-tRNA reductase; aminolevulinic acid; 5-aminolevulinic acid; ALA;
XX tetrapyrrolic pigment; chlorophyll; transgenic plant; ss.
XX Oryza sativa.
XX
XX Key Location/Qualifiers
XX CDS 9..264
XX FT /*tag= a
XX FT /product= "Glu-tRNA reductase"
XX FT /transl_except= "(pos: 234..237, aa: Xaa)"
XX FT /note= "Xaa is an unknown amino acid"
XX
XX FN MO200109304-A2.
XX
XX PD 08-FEB-2001.
XX
XX PF 28-JUL-2000; 2000WO-US21008.
XX
XX PR 30-JUL-1999; 99US-0146600.
XX
XX PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX PI Cahoon RE, Gutteridge S, Harvell LT, Rafalski JA, Tao Y, Weng Z;
XX
XX DR WPI; 2001-159865/16.
XX
XX DR P-PSDB; AAB31961.
XX
XX
XX New nucleic acids encoding an aminolevulinic acid biosynthetic enzyme
XX useful for producing large amounts of the encoded polypeptides used in
XX screening compounds for potential herbicidal activity -
XX
XX Claim 2; Page 51; 77pp; English.
XX
XX The present sequence encodes a Glu-tRNA reductase, which is an
XX aminolevulinic acid biosynthetic enzyme. The enzyme converts
XX Glu-tRNA-Glu to glutamate 1-semialdehyde (GSA) with the concomitant
XX release of tRNA-Glu. GSA aminotransferase then converts GSA to
XX 5-aminolevulinic acid (ALA). ALA is used in the biosynthesis of
XX tetrapyrrolic pigments such as chlorophyll. Nucleic acids encoding
XX aminolevulinic acid biosynthetic enzymes are useful for isolating cDNAs

```

CC and genes encoding homologous proteins from the same or other plant
 CC species, for creating transgenic plants in which the polypeptides are
 CC present at higher or lower levels than normal or in cell types or
 CC developmental stages in which they are not normally found, for
 CC overexpression in bacterial or yeast hosts to efficiently produce
 CC large amounts of the encoded polypeptides which could then be used for
 CC screening different compounds for potential herbicidal activity, and
 CC as hybridisation probes and amplification primers.

XX
 SQ Sequence 510 BP; 144 A; 102 C; 130 G; 111 T; 23 other;

Query Match 11.6%; Score 223.2; DB 22; Length 510;
 Best Local Similarity 70.7%; Pred. No. 1.1e-38;
 Matches 367; Conservative 0; Mismatches 136; Indels 16; Gaps 6;

QY 217 TGGAGGCCCGCAGCGCGTGGCAAGGCTGCAGCGTCCGCCCTCGAGCAGTTCA 276
 DB 1 TGGTATCCCGCAGCGCGTGGCAAGGCTGCAGCGTCCGCCCTCGAGCAGTTCA 60

QY 277 AGATATCCGCCCGCAGCGGTACATGAAGGAGGAGTACCATAGTCTGTAGTGGCCTCAGTG 336
 DB 61 AGATCTCCGCCCGCAGCGGTACATGAAGGAGGAGTACCATAGTCTGTAGTGGCCTCAGTG 120

QY 337 TACACAGCAGCAGTGGAGATGCGTGAAGAACTTGTCTTTCAGAGGAACTGTGGCCCC 396
 DB 121 TACACACTGCACAGTGGAGATGCGTGAAGAACTTGTCTTTCAGAGGAACTGTGGCCCC 180

QY 397 GTGCTATTCAAGACTACTAGCCTGAACCATATTGAAGAGGCTGCTGTTCTT-AGTACC 455
 DB 181 GTGCTATTCAAGACTACTAGCCTGAACCATATTGAAGAGGCTGCTGTTCTTAAAGTACC 240

QY 456 TGTATAGAAATGGAATTTATGTGG---TGGCGCTATCATGAACCGTGGTATCAGAGAA 512
 DB 241 TGTATAGAAATGGAATTTATGTGG---TGGCGCTATCATGAACCGTGGTATCAGAGAA 300

QY 513 GTAGTGGAGTGGATGTCGAAGAAAGTGTATTCCTCGCTTCGAGCTTAGGAGCACCTG 572
 DB 301 AGTGTAACTGGATTTCAGAGAAAGTGTATTCCTCGCTTCGAGCTTAGGAGCACCTG 355

QY 573 TTTATCTTTCGAGCAGTGGATGCCACAGCCATCTGTTGAGGTGTAGCTGGCCTTGAC 632
 DB 356 ATCAAGTNCCTTGANATGTATNCANAGCAATCNGTTGAGGNACNCCGGGC---TNN 412

QY 633 TCTTTGGTTCTCGTGAAGGACAAATCTTGTCTCAGGTTAAACAGTTGTGAGGAGTGA 692
 DB 413 ACCTTGGTTCTTGAAGGAGGAAATCTTGNCANNTAANA----ATTCAAAATGGG 468

QY 693 CAGAACAGTGGAGGTTGGGAAAGAACATCGATAGGATG 731
 DB 469 CAAAAAATTGGAGGNTGGAANANANNCATTTGTTNTAAG 507

RESULT 8
 AAF54813
 ID AAF54813 standard; cDNA; 312 BP.
 XX
 AC AAF54813;
 XX
 DT 15-MAY-2001 (first entry)
 XX
 DE Nucleotide sequence of a corn Glu-tRNA reductase enzyme.
 XX
 KW Glu-tRNA reductase; aminolevulinic acid; 5-aminolevulinic acid; ALA;
 XX tetrapyrrolic pigment; chlorophyll; transgenic plant; ss.
 XX
 OS Zea mays.
 XX
 FH Key
 FT CDS
 FT Location/Qualifiers
 FT 2..192
 FT /*tag= a
 FT /product= "Glu-tRNA reductase"
 FT /transl_except= "(pos: 29..31, aa: Xaa)"
 FT /transl_except= "(pos: 137..140, aa: Xaa)"

FT
 FT
 PN
 XX
 PD
 XX
 PF
 XX
 PR
 XX
 PA
 XX
 PI
 XX
 DR
 DR
 XX
 PT
 PT
 PT
 PS
 XX
 CC
 CC
 CC
 CC
 CC
 CC
 CC
 CC
 CC
 CC
 CC
 CC
 CC
 SQ

WO200109304-A2.
 08-FEB-2001.
 28-JUL-2000; 200WO-US21008.
 30-JUL-1999; 99US-0146600.
 (DUPO) DU PONT DE NEMOURS & CO B I.
 Cahoon RE, Gutteridge S, Harvell LT, Rafalski JA, Tao Y, Weng Z;
 WPI; 2001-159865/16.
 P-PSDB; AAB31959.
 New nucleic acids encoding an aminolevulinic acid biosynthetic enzyme
 useful for producing large amounts of the encoded polypeptides used in
 screening compounds for potential herbicidal activity -
 Claim 2; Page 46; 77pp; English.
 The present sequence encodes a Glu-tRNA reductase, which is an
 aminolevulinic acid biosynthetic enzyme. The enzyme converts
 Glu-tRNA-Glu to glutamate 1-semialdehyde (GSA) with the concomitant
 release of tRNA-Glu. GSA aminotransferase then converts GSA to
 5-aminolevulinic acid (ALA). ALA is used in the biosynthesis of
 tetrapyrrolic pigments such as chlorophyll. Nucleic acids encoding
 aminolevulinic acid biosynthetic enzymes are useful for isolating cDNAs
 and genes encoding homologous proteins from the same or other plant
 species, for creating transgenic plants in which the polypeptides are
 present at higher or lower levels than normal or in cell types or
 developmental stages in which they are not normally found, for
 overexpression in bacterial or yeast hosts to efficiently produce
 large amounts of the encoded polypeptides which could then be used for
 screening different compounds for potential herbicidal activity, and
 as hybridisation probes and amplification primers.
 Sequence 312 BP; 79 A; 68 C; 90 G; 69 T; 6 other;

Query Match 9.3%; Score 179.8; DB 22; Length 312;
 Best Local Similarity 95.0%; Pred. No. 2.3e-29;
 Matches 207; Conservative 0; Mismatches 8; Indels 3; Gaps 2;

QY 224 CCAGCGCAGCGCGCTGGCAAGGCTGCCAGGCTGCCGCCCTCGAGCAGTTCAAGATATC 283
 DB 1 CCAGCGCAGCGCGCTGGCAAGGCTGCCAGGCTGCCGCCCTCGAGCAGTTCAAGATATC 60

QY 284 CGCCGACCGGTACATGAAGGAAAGAGTACCATAGTGTGTAGTGGCTCAGTGTACACAC 343
 DB 61 CGCCGACCGGTACATGAAGGAAAGAGTACCATAGTGTGTAGTGGCTCAGTGTACACAC 120

QY 344 AGCACCACTGGAGATG--CGTGAAGAACTTGTCTTTCAGAGGAACTGTGGCCCCGTCT 401
 DB 121 AGCACCACTGGAGATGCGGTGTAAGAACTTGTCTTTCAGAGGAACTGTGGCCCCGTCT 180

QY 402 ATTCAGAAAC--TCACTAGCCTGAACCATATTGAAGAGG 438
 DB 181 ATTCAGAACTTTACTAGCCTGAACCATATTGAAGAGG 218

RESULT 9
 ABX21149
 ID ABX21149 standard; cDNA; 384 BP.
 XX
 AC ABX21149;
 XX
 DT 10-FEB-2003 (first entry)
 XX
 DE Human GDP-mannose 4,6-dehydratase (GM4, 6D) DNA #3206.
 XX

KW Human; GDP-mannose 4,6-dehydratase; GM4,6D; gene; ss; inflammation;
 KW cellular fucosylation; glycoconjugate fucosylation; transplant rejection;
 KW arthritis; asthma; sepsis; reperfusion injury; stroke; infection;
 KW complex carbohydrate; gene replacement therapy; immunosuppressive;
 KW antiinflammatory; antiarthritic; antibacterial; cerebroprotective;
 KW antiasthmatic; vasotropic.
 XX
 OS Homo sapiens.
 XX
 PN US2002110548-A1.
 XX
 PD 15-AUG-2002.
 XX
 PP 11-JUN-2001; 2001US-0878574.
 XX
 PR 22-NOV-1996; 96US-0753233.
 PR 03-DEC-1997; 97US-0984246.
 PR 09-SEP-1998; 98US-0149674.
 PR 14-JUN-1999; 99US-0333177.
 XX
 XX (GEMY) GENETICS INST INC.
 PA Sullivan F, Kriz R, Kumar R;
 PI WPI; 2003-0666673/06.
 XX
 XX New composition comprising GDP-mannose 4,6-dehydratase (GM4,6D)
 PT peptide, for manufacturing complex carbohydrates, or as targets for
 PT screening GM4,6D antagonists for treating e.g. arthritis, or transplant
 PT rejection -
 XX
 PS Disclosure; SEQ ID NO 3208; 6pp; English.
 XX
 XX The invention relates to a composition comprising a human GDP-mannose
 CC 4,6-dehydratase (GM4,6D) peptide. The peptide is useful for identifying
 CC GM4,6D inhibitors. GM4,6D inhibitors are useful for reducing inflammation
 CC in a mammalian subject and for treating or ameliorating diseases affected
 CC by the level of cellular fucosylation or diseases affected by the
 CC fucosylation of glycoconjugates. These diseases include arthritis,
 CC transplant rejection, asthma, sepsis, reperfusion injury, stroke or
 CC infection. The GM4,6D peptide or a polynucleotide encoding it is also
 CC useful for manufacturing complex carbohydrates and as targets for
 CC screening small molecule antagonists of the activity of the enzyme. The
 CC polynucleotide is useful in developing an assay for defects in the
 CC enzyme, as well as in gene replacement therapy. Sequences
 CC ABX17942-ABX17944 and ABX17947-ABX33716 represent DNA molecules encoding
 CC human GM4,6D peptides of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format directly from USPTO
 CC at seqdata.uspto.gov/sequence.html.
 XX
 XX Sequence 384 BP; 128 A; 61 C; 95 G; 100 T; 0 other;
 XX
 Query Match 6.5%; Score 125.6; DB 25; Length 384;
 Best Local Similarity 64.6%; Pred. No. 1.3e-17;
 Matches 204; Conservative 0; Mismatches 109; Indels 3; Gaps 1;
 QY 1366 AGAAGCTGAGTGGCTACCGGACAGATCAGGAGCCATCGAGGAGCTGAGCACCG 1485
 DB 1 AGAATTGAGGCTTATGCTGAAAGAAATCAGGCTGCTGAGCTTGAAGAGTCTTAGGTA 60
 QY 1426 AAGTAGGTGAGGACGCGCTTACCAAGAGATGAGGAGAGCCATCGAGGAGCTGAGCACCG 1485
 DB 61 AGATGGTGTATATATAC---CAAGAAACCGGAGAGCTGTGATGACCTTAGTCGGG 117
 QY 1486 GCATGCTTAAAGCTCTCCATGCGCCCTCGACACCTGAGGTGCGACGCGACGACA 1545
 DB 118 GTATAGTGAATAAGTTGCTTATGTTCCAAATGCAACATTAAAGGTGTGATGGAACGACA 177
 QY 1546 GCCGACCCCTTACGAGAGCTTCGAGACATCGACGCTTCAACCGGATGTTACGCTCG 1605
 DB 178 GCCGGACTTCTAGTGAGACACTGGAGAACATGAATGCTTTGAATAGGATGTTCAACCTTG 237

QY 1606 ACATGGAGAAGCGGATCATCGAGCAGGAAGATCAAGGCCAGGTGAGAGACACAAAAT 1665
 DB 238 AGACAGAAATATCTGTTTTGGAGGAGGAAGATTCGAGCAAGGTCGACAAAACCGAAT 297
 QY 1666 GAGGCCAGGAACAAT 1681
 DB 298 GAAATCTAACACCAAT 313
 RESULT 10
 ABX23406
 ID ABX23406 standard; CDNA; 275 BP.
 XX
 AC ABX23406;
 XX
 DT 10-FEB-2003 (first entry)
 XX
 DE Human GDP-mannose 4,6-dehydratase (GM4,6D) DNA #5463.
 XX
 KW Human; GDP-mannose 4,6-dehydratase; GM4,6D; gene; ss; inflammation;
 KW cellular fucosylation; glycoconjugate fucosylation; transplant rejection;
 KW arthritis; asthma; sepsis; reperfusion injury; stroke; infection;
 KW complex carbohydrate; gene replacement therapy; immunosuppressive;
 KW antiinflammatory; antiarthritic; antibacterial; cerebroprotective;
 KW antiasthmatic; vasotropic.
 XX
 OS Homo sapiens.
 XX
 PN US2002110548-A1.
 XX
 PD 15-AUG-2002.
 XX
 PF 11-JUN-2001; 2001US-0878574.
 XX
 PR 22-NOV-1996; 96US-0753233.
 PR 03-DEC-1997; 97US-0984246.
 PR 09-SEP-1998; 98US-0149674.
 PR 14-JUN-1999; 99US-0333177.
 XX
 XX (GEMY) GENETICS INST INC.
 PA Sullivan F, Kriz R, Kumar R;
 PI WPI; 2003-0666673/06.
 XX
 XX New composition comprising GDP-mannose 4,6-dehydratase (GM4,6D)
 PT peptide, for manufacturing complex carbohydrates, or as targets for
 PT screening GM4,6D antagonists for treating e.g. arthritis, or transplant
 PT rejection -
 XX
 PS Disclosure; SEQ ID NO 5465; 6pp; English.
 XX
 XX The invention relates to a composition comprising a human GDP-mannose
 CC 4,6-dehydratase (GM4,6D) peptide. The peptide is useful for identifying
 CC GM4,6D inhibitors. GM4,6D inhibitors are useful for reducing inflammation
 CC in a mammalian subject and for treating or ameliorating diseases affected
 CC by the level of cellular fucosylation or diseases affected by the
 CC fucosylation of glycoconjugates. These diseases include arthritis,
 CC transplant rejection, asthma, sepsis, reperfusion injury, stroke or
 CC infection. The GM4,6D peptide or a polynucleotide encoding it is also
 CC useful for manufacturing complex carbohydrates and as targets for
 CC screening small molecule antagonists of the activity of the enzyme. The
 CC polynucleotide is useful in developing an assay for defects in the
 CC enzyme, as well as in gene replacement therapy. Sequences
 CC ABX17942-ABX17944 and ABX17947-ABX33716 represent DNA molecules encoding
 CC human GM4,6D peptides of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format directly from USPTO
 CC at seqdata.uspto.gov/sequence.html.
 XX
 XX Sequence 275 BP; 60 A; 51 C; 84 G; 80 T; 0 other;
 XX
 Query Match 6.4%; Score 122.6; DB 25; Length 275;

Db 1178 CGCTATGTGGAACCTTTGCAAGCCAGGCAAGCGGCGCTCGCGCC 1237

Qy 1380 TACGCGCAGGATCAGGGCTCGAGCTCGAGAAGTGCTTCAGAA 1426

Db 1238 TTTGGCATGACACCCGCGAGCTGTGSCCAAGGCGGCCAGCA 1284

RESULT 12

RESOL 12
ABK73683
ID ABK73683 standard; DNA; 1360 BP.

AC ABK73683;
XX
DT 13-AUG-2002 (first entry)

DE Bacillus licheniformis genomic sequence tag (GST) #974.

Differential gene expression; genomic sequenced tag; GST;
KW
KW
KW
KW
altered culture condition; environmental stress;
physiological provocation; ds.

AA
OS
Bacillus licheniformis.

XX PN WO200229113-A2.

XX
PD
11-APR-2002-XX
PF 05-OCT-2001: 2001WO-US31437.XX
PR 06-OCT-2000; 2000US-0680598

PR 27-MAR-2001; 2001US-279526P

PA (NOVO) NOVOZYMES BIOTECH INC.

PA (NOVO) NOVOZYMES AS.

PI Berka R, Clausen IG;

DR WPI; 2002-416684/44.

Monitoring differential expression of several genes in first *Bacillus* cell relative to expression of same genes in one or more second *Bacillus* cells, by using substrate containing *Bacillus* genomic sequenced tag array -

PS Claim 4: SEO ID NO 974: 200pp: English.

The invention describes a method of monitoring differential expression of genes in a first *Bacillus* cell relative to expression of the genes in other *Bacillus* cells, comprising hybridising labelled nucleic acid probes isolated from *Bacillus* cells to a substrate containing array of *Bacillus* genomic sequenced tags (GST), examining the array, and determining relative gene expression by an observed hybridisation reporter signal of a spot in the array. The method is useful for measuring the expression of genes in a first *Bacillus* cell relative to expression of the same genes in one or more second *Bacillus* cells. The method is useful for monitoring global expression of several genes from a *Bacillus* cell, discovering new genes, identifying possible functions of unknown open reading frames and monitoring gene copy number variation and stability. Monitoring changes in expression of genes may be used to provide a representation of the way in which *Bacillus* cells adapt to changes in culture conditions, environmental stress or other physiological provocation. Extensive follow-up characterisation is unnecessary, when one spot on an array equals one gene or one open reading frame, since sequence information is available. This sequence represents a genomic sequence tag (GST) used in the method of the invention.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at

SQ Sequence 1360 BP; 367 A; 293 C; 374 G; 326 T; 0 other;

Query Match

4.5%; Score 86.6; DB 24; Length 1360;

Best Local Similarity 44.7%; Pred. No. 5.7e-09; Matches 485; Conservative 0; Mismatches 579; Indels 21; Gaps 3;	
Qy	315 ATAGCTGTGATAGCGCTCAGTGTACACACAGCACACAGTGGAGATGCGTGAAGAACTTGCT 374
Dd	7 ATACTGTGCTGGATTGGATTACAATCAGCCCTGTTGTGATATAGTGAAGATTAAAC 66
Qy	375 GTTGCAGAGAACTGTGGCCCGCTGCTATTCAAGAACTCACTAGCCCTGAACCACTATTGA 434
Dd	67 TTTTCAGCCGATGAGCTGGGCGAGCGATGGCGCAATTGAAGAAGAAAAAGCATACTT 126
Qy	435 GAGGCTGCTGTTCTTAGTACTGTTATAGAAATGGAATTTATGTGTGGCGCTATCATGG 494
Dd	127 GAGAAACATCATCGTGTGCAATGCAACCGGACAGAGCTGTATGCAGTCTGTGATCAGCTG 186
Qy	495 AACCGTGTATCAGAGAAGTGTGGATGTCGAAGAAAAAGTGTATTCCTCCGCTTCC 554
Dd	187 CACACCGGAGCTTATTATATGAATAATGTTTTTGGCGAATGCGTTCGATTGAACAAAGAG 246
Qy	555 GAGCTTAGGAGCACCTGTTTCATCTTTCGAAGCAGTGTGTCACAGCCATCTGTTTGGAG 614
Dd	247 GACATTTTCGCTTACTTGAAGTTTTATGAAAAAGCGCGGCGCTTGAACACCTGTTCCCG 306
Qy	615 GTGTCAGCTGGCTTGACTCTTTGGTCTCGGTGAAGGACAAATCTTGTCTCAGGTTAAA 674
Dd	307 GTATCGTGCAGATTGGAATTCGATTCGATTTGAGAGAAACGAGATTTTGGGCCAAGTCGT 366
Qy	675 CAAGTTGTGAGGAGTGCAGACAGTGTGGAGGCTTCGGAAGAAAGAACATCGATAGGATGTT 734
Dd	367 ACAAGTTTTAAGCTGGCTCAGGAAGAGAGACAATTTGGTACCGTTTTCAACTATTATGTTT 426
Qy	735 AAGGATGCAATCACTGCTGGAAGACGCTGTCGCGAGCGAGACCAACATATCATCTGCTGCT 794
Dd	427 AAGCAGCGCTCACGGTTCGCAAGCGCTGCCATGCGGAACGGACATCTCTTCAATGCT 486
Qy	795 GTTTCGTGCTAGTTACGGGGGGTTGAACCTGGCCCTGATGAAGCTTCGGAAGTCTGAAGCA 854
Dd	487 GTATCCGCTCAGCTATGCGCGCGTCGAGCTTGCAGGAAAAATCTTCGCGCACTTCCGAT 546
Qy	855 CTGTGAGCTAGGATGCTTCTGATGTGCTGTGTAAATGGGAAGCTAGTGATCAACAT 914
Dd	547 AAACACG-----TGCTGATCTTCGGGGCGGCAAAATGGGTGAGCTTCGGGTTCAAAAC 600
Qy	915 CTGTTTGCCAAAGGATGCAAGAAAGTTGTTGTGGTCAACCGCTCCGTGGAAGGGTGGAT 974
Dd	601 CTCACGGCCACGGAATCGGACAGGTTACGGTATCAACAGAACATTTTCAAGCGAAG 660
Qy	975 GCTATTGTCAGGAGATGAAGATATAGAGATGCTGTACAGGCCTCTCTCAGACATGTAT 1034
Dd	661 GAGCTTGCCGCGCTTTTTCAGCGCGGGGAAAGCTTGAA-----TGAGCTGCAA 711
Qy	1035 CAAGCTGCTCTCAAGCTGATGCTGTTTCAACGACCGCATCTGAACACTTCATTGTTTC 1094
Dd	712 TGGCTTTGATGAAGAGCCGATATTTCTCATGATTGACCGCGCGCAGCGGCTATGTCGT 771
Qy	1095 GCAAAAGAACACGACAGGACATCCCCCTGCTCTGTGATCTATGGAAGGTGTTCCGCTG 1154
Dd	772 ACAAGAGATGATCGAACATGTCAATAAGCT-----TCGCAAGGCTGTCCGCTTTT 825
Qy	1155 TTTGTGCAATATCTCTCTCCAGGAATGTACGGCATGTGTGTCTGAAGTTGCGCTGCA 1214
Dd	826 ATGGTGGACATTTCCGCTGCGGGGATCTTGACCCGGCTCTCAGCGAAGTGAAGAGCGGTC 885
Qy	1215 CGAGTGTACAAATGTCACGACCTTGAAGAGGTTGGTGAAGCCCAAGAGAGGACCGGCTC 1274
Dd	886 TTCCTGTACGATATTGACCATCTTGAAGGGATGCTGCAAGAAAAATTTGAAGAAACGCCAG 945
Qy	1275 AGGAAAGCAATGAGGCGCAGACAATCATCAACGAAGAACTGAGACGGTTCCGAGCATGG 1334
Dd	946 GCGGTTGCCGAGAGGTGGGCTGATTATTGAAGCTGAGATCGTTCTTTTAAACAAATGG 1005
Qy	1335 AGGGAATCGCTGGAGACCGTTCCGACCATCAAGAAAGCTGAGTCTGTAAGCGGACAGGATC 1394

Best Local Similarity 46.9%; Pred. No. 3.2e-08;
Matches 380; Conservative 0; Mismatches 410; Indels 21; Gaps 3;
Qy 555 GAGCTTAGGAGGACCTGTTCATCTTGGGAGCAGTGATGCCACAGCCCATCTGTTTGAG 614

Qy 615 GGTGCACTGGCCTTGACTCTTTGGTCTCGTGAAGGACAATCTTGCTCAGGTTAA 674

Db 107 GTTGCTGTGGTTTAGACTCTATGGTGATCGCGAAACGCAATCTTTGGACAAGTCAAG 166

QY 675 CAAGTTGTGAGAGTGA CAGAACACGTGGAGCCTCGGGAAGAAACATCGATAGGATGTTT 734

Dh 167 CAGGCATTTTTCTCTGCC CAGAAGAAAGTTCACAGGCATCTGTGTTTAAACAGCTATATTT 226

735 AAGGATGCAATCACTGGCTGGAAAGCGGTGTCCGCAGCGAGCCAACATATCATCTGGTGCT 794

Db	227	AAACAAGCGGTACGCTCGGCAACGCTGCACTCGTCAACGGAAATTTCTTCACAGCC	286
Ov	795	GTTTCTCTCAGTTTCAGCGGCGGTTGAACATCGCCCCCATGAAGCTTCGGAAGTCTCAAGCA	854

287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243 1244 1245 1246 1247 12

[illegible]

db

347 AACCACGCTC-----TCCCTTTATTTTTTCCCCCCCCTTCCTCAAAATTCTCCCGCCAGCAT

915 CTGGTTGCCA AAGGATGCA AGA AGGTTGTCTGCTCA A CCGCTCCCTGCA A CCGCTGCAT 974

db
401 CTTTATTTCTGAAATGGGGGCTGCTTCCATTATGCTCTCATCTCCGACAAACACAAACACCCCTC 460

975 GCTATTCTGTCAGGAGATGAAAGATATAGAGATCGTGTACAGGCCTCTCTCAGACATCTAT 1034

Db 461 GAAC TGGCAAGCCAGTTTTC TGG ACGCT --- CCTCTTTTCCTCAATTCAT 511

1035 CAAGCTGCTGAAGCTGATGTCGTGTCCAGCACTCTGAAACCTTCATTGTTTC 1084

db 512 GATGCTTAAAAGAGCGAGTGGTGTGCTTACCTCCCGCCGCGCTTCTCTTC 571

QY 1095 GCAAAAGAACACGACGAGGCACCTCCCCCTGTCCTCTGATACTATGGGAGGTGTTGCCCTG 1154

Db 572 AAAAAGGAAATGCCCGAGTGGCC-----TTAAAAACGGAAGCGAGGCCGCTGTTT 625

QY 1155 TTGTTCGACATATCTGTCCCCCAGGAATGTTCAGCGCATGTGTCTGAAGTTGGCGCTGCA 1214

Db	626	GCAGTTTGATATCGAGTGCACCGGGATATTGACCCCGGAAATTGCACAAATTAGCGACGTT	685
Qy	1215	CGAGTGTACAATGCTCGACGACTTTGAAGAGGTGGTGGGAAGCCCAACAAGCAGGACCGGCTC	1274
Db	686	TACTTATACGACATTGATGACTTACAGACANTGTTTGAAGCCCATATAAAGAAGCGCAAA	745
Qy	1275	AGGAAGCAATAGGAGCGGCAGACAAATCATCAACGGAAGAACTGAGACGGTTCGAGGCGATCG	1334
Db	746	AAAGAGCGTCAAAAAATCGTGTATTGATTTGAGGCAGAGATCGATGACTTTAAAGCATCG	805
Qy	1335	AGGGACTCGCTGGAGACCGTTCCGACCATCA	1365
Db	806	CTAAACACACTAGGGTGGTGCCTCTTTATTA	836

RESULT 14
AAZ54023
ID AAZ54023 standard; DNA; 1248 BP.
XX
AC AAZ54023.

XX DE NEIGER'S MENTIONED IN THE 1900'S

Not a member of the

KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy; do

XX OS *Neisseria meningitidis*.
 XX PN W09957280-A2.
 XX PD 11-NOV-1999.
 XX PF 30-APR-1999; 99WO-US09346.
 XX PR 01-MAY-1998; 98US-0083758.
 XX PR 31-JUL-1998; 98US-0094869.
 XX PR 02-SEP-1998; 98US-0098994.
 XX PR 02-SEP-1998; 98US-0099062.
 XX PR 09-OCT-1998; 98US-0103749.
 XX PR 09-OCT-1998; 98US-0103794.
 XX PR 09-OCT-1998; 98US-0103796.
 XX PR 25-FEB-1999; 99US-0121528.
 XX PA (CHIR) CHIRON CORP.
 XX PA (GENO-) INST GENOMIC RES.
 XX PI Fraser C, Galeotti C, Grandi G, Hickey E, Massignani V, Mora M;
 XX PI Pecersen J, Pizzo M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 XX PI Tettelin H, Venter JC;
 XX DR WPI; 2000-062150/05.
 XX DR P-PSDB; AAY75261.
 XX PT Novel *Neisseria* polypeptides predicted to be useful antigens for
 XX PT vaccines and diagnostics -
 XX PS Claim 7; Page 997; 1453pp; English.
 XX CC AA253015 to AA254536, AA254577 to AA254615, and AAY74253 to AAY75941
 CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC *Neisseria* bacteria (e.g. meningitis and septicemia), to detect the
 CC presence of *Neisseria* bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.
 XX SQ Sequence 1248 BP; 300 A; 305 C; 363 G; 280 T; 0 other;

Query Match 4.1%; Score 78.6; DB 21; Length 1248;
 Best Local Similarity 47.5%; Pred. No. 3e-07;
 Matches 317; Conservative 0; Mismatches 329; Indels 21; Gaps 2;

QY 325 TAGCCCTCAGTGATACACAGCAGCCAGTGGAGATGCGTGAAACCTTGCTGTTCCAGAGG 384
 Db 17 TCGGACTCAATCATCAACCCGACCTTTAAGCATACCGGAAAGCTGGCGTTTCCCGCG 76
 QY 385 AACCTGTGGCCCGTGCTATTCAAGAACTCACTAGCCTGAACCATATTGAAGAGGCTG 444
 Db 77 CCTGCTGCCGGAAGCCGTCGCAATCTTGCCGAAGCAATGCGGCAACGAGGCGGTAA 136
 QY 445 TTCTTAGTACTGTAATAGAAATATATGTTGGTGGCGCTATCATGAAACCGTGGTA 504
 Db 137 TCCTTTCTACTGCAACCGTACCGAGCTTTACTGTGTAG-----GTGATT 181
 QY 505 TCAGAGAGTAGTGGACTGATGTCGAAGAAAGTGTATTCCGGCTTCGAGCTTAGG 564
 Db 182 CGGAAGAAATCATCCGTTGGCTGCGACATATCACAGCCTTCCCATAGAAAGAAATCAGCC 241
 QY 565 AGCACCTGTTTCATCTTCGGAAGCAGTGTATGTCACACCGCCATCTGTTTGGTGTGAC 624
 Db 242 CCTACCTTTATATCTTTGGGATGACGAGACTGTGCCCATGCTTTCCGCGTCCCTGCG 301

QY 625 GCCTTGACTCTTTGGTTCTCGTGAAGGACAAATCCTTGCTCAGGTTAAACAAGTTGTA 684
 Db 302 GCTTGGATTCGATGGTGTGGGCCAGCGCGCAGATTTTAGGACAGATTAAAGGATCGGTCA 361
 QY 685 GGAGTGGACAGAAACAGTGGAGGCTTGGGAAAGAAACATCGATAGGATGTTCAAGGATGCAA 744
 Db 362 GGGTTGCTCAAGACGAGAAAGTATGGTAAGAACTCAATGCCCTGTTCCAAAAAACCCT 421
 QY 745 TCATGCTGGAACCGGTGTCGACGAGACCAACATATCATCTGGTGTCTGTTCTGTGTA 804
 Db 422 TTTCTGTTGCTAAAGAGTCCGTACCGATACTGCGCGGCGGAAACCTCGGTTTCCATGG 481
 QY 805 GTTCAGCGCGGTTGAACTGCGCCCTGATGAAGCTTCCGAAGCTGGAAGCACTGTGAGCTA 864
 Db 482 CTTCCGCTTCGTCGTAAGTTGGCAGAGCAGATTTTCCCGACATCGGCGATTTGAATG--- 538
 QY 865 GGATGCTTCTGATTGGTGGTAAATGGAAAGCTAGTATCAAAACATCTGGTTGCA 924
 Db 539 ---TCTTGTATTATCGGTGCGGTGAGATGATTGAGCTGGTTGCCACTTATTTTCCGCCA 595
 QY 925 AAGATGCAAGAGGTTGTTGTGTGAACCGCTCCGTGGAAAGGGTGGATGCTATTCTGTG 984
 Db 596 AAAGTCCCGGCTGATGACGTTTGCCAAACCGAGCTGGCGGTGTCACAGGAGTTGTGCG 655
 QY 985 AGGAGAT 991
 Db 656 ACAAGCT 662

RESULT 15
 ABQ67195
 ID ABQ67195 standard; DNA; 495269 BP.
 XX AC ABQ67195;
 XX DT 29-AUG-2002 (first entry)
 XX DE *Listeria innocua* contig DNA sequence #8.
 XX KW Antibacterial; *Listeria*; food contamination; mutational analysis;
 XX KW infection; ds.
 XX OS *Listeria innocua*.
 XX PN W0200228891-A2.
 XX PD 11-APR-2002.
 XX PF 04-OCT-2001; 2001WO-FR03061.
 XX PR 04-OCT-2000; 2000FR-0012697.
 XX PA (INSP) INST PASTEUR.
 XX PA (CNRS) CNRS CENT NAT RECH SCI.
 XX PI Kunst F, Glaser P;
 XX DR WPI; 2002-332479/37.
 XX PT New genomic sequences from *Listeria* species, useful for detection,
 PT treatment and prevention of infection, also related polypeptides,
 PT antibodies and modulators -
 XX PS Claim 5; SEQ ID 8; 180pp; French.
 XX CC The present invention relates to nucleic acid sequences
 CC (ABQ67188-ABQ671212) from *Listeria* sp. The sequences are useful as probes
 CC and primers for identification and/or detection of *Listeria* (e.g. as
 CC contaminants in foods, or mutational analysis) and for analysis of
 CC gene expression. Proteins encoded by the nucleic acid sequences can be
 CC used to screen for compounds that modulate gene expression, replication
 CC and pathogenicity of *Listeria* (potential therapeutic agents), also for
 CC treating infections by *Listeria*, and are useful as immunogens in

CC anti-Listeria vaccines.

CC anti-Distoceria vaccines.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published pct sequences.

XX
XX

WY

SQ Sequence 495269 BP; 159529 A; 88849 C; 93135 G; 153754 T; 2 other;

TABLE 7. EFFECT OF CURE IN CEMENT LINING ON CORROSION RATE

Quinn, M. J.

Query Match 3.9%; Score 75.4; DB 24; Length 495269;

Best Local Similarity 46.1%; Pred. No. 1.1e-05;

2000 Local Similarity 40.1%, Read: 10. 1.10-0.0,
 Matches 203: Conservative 0. Mismatches 226. Indels 1

Matches 293; Conservative 0; Mismatches 336; Indels 6; Gaps 1;

324 ATAGGCTCAGTGTACACACAGCACCGTGGATGGTGAAAACCTTCCTGCTCCACAC

324 ATAGGCCCTCAGTGTACACACAGACCTGCTGTGGCAGAG 383

Db 160090 ATGGGGCTGAATCATCACAGCACCAATCGACATCCGGGAAAAATTAGTTTAAAGAA 160149

52 TGGGCTGATCATCAGCAGCAGCCATCGACATCCGCGAATAAAATAGTIIITAAAGAA TGGTATG

QY 384 GAACTGTGGCCCGTGCTATTCAAGAACTCACTAGCCTGAACCATATTGAAGAGGCTGCT 443

[illegible]

Db 160150 ACGGAAGAAGAAATGGCCTTAGTAACATTACTGCAGGAAAAAAGTATCCTCGAAAAATGTT 160209

C0Z0N0Y T7D1XK5H0C1CCYU76UEEEEEEEOIIC0T0OIT7EVEIOOOOTTEWIEEIOOI IEEEE --

[illegible]

QY 444 GTTCTTAGTACCTGTAATAGAAATTTATGTGGTGGCGCTATCATGGAACCGTGGT 503

A vertical ruler with markings from 0 to 12 inches. The markings are in inches, with major ticks every inch and minor ticks every half inch. The numbers 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, and 12 are printed along the right side of the ruler.

150318

D5 160210 ATTATTTCGACATGCAATCGTACGGAGATTGTGGCAGTTGTGCATCAAATACATACAGGC 160269

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841

[illegible]

504 ATCAGAGAAGTAGTGGACTGGATGTCGAAGAAAGTGGTATTCCCGCTTCCGAGCTTAGG 563

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841

Search completed: D

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 12:57:36 ; Search time 1512 Seconds
(without alignments)
4229.236 Million cell updates/sec

Title: US-10-018-902-3

Perfect score: 1924

Sequence: 1 ccacgcgcgcgcacataa.....aaaaaaaaaaaaaaaaaag 1924

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:**
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US05_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	679.6	35.3	1632	10	US-09-938-842A-1201
2	125.6	6.5	384	10	US-09-878-574-3208
3	122.6	6.4	275	10	US-09-878-574-5465
4	96.6	4.5	1360	10	US-09-974-300-974
5	83	4.3	1131	10	US-09-974-300-5501
6	68	3.5	271	10	US-09-878-574-6807
7	60.6	3.1	1707	15	US-10-156-761-4723
8	60.6	3.1	9025608	15	US-10-156-761-1
9	59.2	3.1	1496	8	US-08-781-986A-274
10	56.2	2.9	380	9	US-09-770-791-195
11	54.4	2.8	1389	10	US-09-738-626-470
12	54.4	2.8	3309400	10	US-09-738-626-1
13	51	2.7	3303	15	US-10-156-761-5384
14	51	2.7	9025608	15	US-10-156-761-1
15	50.6	2.6	1965	13	US-10-190-435-50

16	50.6	2.6	1965	13	US-10-190-305A-55	Sequence 55, Appl
17	50.6	2.6	1971	13	US-10-190-435-49	Sequence 49, Appl
18	50.6	2.6	1971	13	US-10-190-305A-54	Sequence 54, Appl
19	50.6	2.6	2445	13	US-10-190-435-43	Sequence 43, Appl
20	50.6	2.6	2445	13	US-10-190-305A-37	Sequence 37, Appl
21	50.6	2.6	2457	13	US-10-190-435-44	Sequence 44, Appl
22	50.6	2.6	2457	13	US-10-190-435-45	Sequence 45, Appl
23	50.6	2.6	2457	13	US-10-190-305A-38	Sequence 38, Appl
24	50.6	2.6	2457	13	US-10-190-305A-39	Sequence 39, Appl
25	50.6	2.6	3162	13	US-10-190-435-18	Sequence 18, Appl
26	50.6	2.6	3231	13	US-10-190-435-60	Sequence 60, Appl
27	50.6	2.6	3231	13	US-10-190-305A-84	Sequence 84, Appl
28	50.6	2.6	3234	13	US-10-190-435-51	Sequence 51, Appl
29	50.6	2.6	3234	13	US-10-190-305A-56	Sequence 56, Appl
30	50.6	2.6	3462	13	US-10-190-435-16	Sequence 16, Appl
31	50.6	2.6	3531	13	US-10-190-435-13	Sequence 13, Appl
32	50.6	2.6	3537	13	US-10-190-435-14	Sequence 14, Appl
33	50.6	2.6	3537	13	US-10-190-435-15	Sequence 15, Appl
34	50.6	2.6	3597	13	US-10-190-435-46	Sequence 46, Appl
35	50.6	2.6	3597	13	US-10-190-305A-40	Sequence 40, Appl
36	50.6	2.6	3607	13	US-10-190-435-48	Sequence 48, Appl
37	50.6	2.6	3607	13	US-10-190-305A-42	Sequence 42, Appl
38	50.6	2.6	3624	13	US-10-190-435-47	Sequence 47, Appl
39	50.6	2.6	3624	13	US-10-190-305A-41	Sequence 41, Appl
40	50.6	2.6	3930	13	US-10-190-435-9	Sequence 9, Appl
41	50.6	2.6	3930	13	US-10-190-435-10	Sequence 10, Appl
42	50.6	2.6	3930	13	US-10-190-435-11	Sequence 11, Appl
43	50.6	2.6	4419	13	US-10-190-435-19	Sequence 19, Appl
44	50.6	2.6	4419	13	US-10-190-305A-14	Sequence 14, Appl
45	50.6	2.6	4713	13	US-10-190-435-59	Sequence 59, Appl

ALIGNMENTS

RESULT 1

US-09-938-842A-1201
; Sequence 1201, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1201
; LENGTH: 1632
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1201

Query Match	35.3%	Score	679.6;	DB 10;	Length	1632;
Best Local Similarity	68.6%;	Pred. No.	8.2e-177;			
Matches	985;	Conservative	0;	Mismatches	439;	Indels 12; Gaps 3;
Qy	235	CCGTGGCAGAGGCTGCCAGGCTGCCGCCCTCGAGCAGTTCAAGATAT---	CCGCCGACC	291		
Db	206	CTGGCTTAATGTCAGTAGCATCTCTGCTCTTGAACTCAAGAACTCTCAGCTGATC	265			
Qy	292	GGTACATGAAGGAAGGAGTACCATAGCTGTGATAGGCTTCAGTGTTACACACACACAG	351			
Db	266	GATATACAGGAAGGAAGGAGTATTTGTTGATGGACTTAGTATTACACACAGCTCTG	325			

QY 352 TGGAGATCGGTGAAAACTTGTCTGTTCAGAGGAACCTGTGGCCCGTGTCTATTCAAGAAC 411
DB |||||
QY 356 TTGAGATCGGTGAGAAGCTTGTCTATACCTGAGCTGGAATGGCCAAAGAGCTATTGCTGAAT 385
DB |||||
QY 412 TCACCTAGCTGAACCATATTGAAGAGGCTGTCTTTAGTACCTGTGTAATAGAAAGGAAA 471
DB |||||
QY 386 TGTGTGGTTGNAATCATATTGAGAGAGCTGTCTGTCTTAGTACTTGTAAACCGTATGGAGA 445
DB |||||
QY 472 TTTATGTGGTGGCGCTCATATGGAACCGTGGTATCAGAGAAAGTAGTGGACTGGATGTGCA 531
DB |||||
QY 446 TTTATGTTTTAGCTTTATCTCAGCATCGTGGAGTTTAAAGAGTTTACTGAATGGATGTCAA 505
DB |||||
QY 532 AGAAAAGTGGTATTTCCGCTTCAGAGCTTAGGGAGCACTGTTCATCTTGGCAAGCAGTG 591
DB |||||
QY 506 AGACGAGTGGGATCCCGATTTCAGAGATTGTTCACACCGCTTTCTGTTGTATCAACAAGG 565
DB |||||
QY 592 ATGCCACACCGCATCTGTTTGAAGTGTCTAGCTGGCGCTTGACTCTTTTGGTTCTCGGTGAAG 651
DB |||||
QY 566 ATGCTTACACAGCATATATTGAGTCTCAGCTGGTCTTGACTCTCTTGTCTTAGAGAGAG 625
DB |||||
QY 652 GACAAATCTTGTCTCAGGTTAAACAAAGTTGTGAGAGTGGACAGAACAGTGGAGCTTGG 711
DB |||||
QY 626 GTCAGATCTTGCACAGAGTGAACAAAGTTGTGAAAGTTGGTCAAGGAGTGAATGGCTTTG 685
DB |||||
QY 712 GAAAGACATCGATAGGATGTTCAAGGATGCAATCACTGCTGGAAGCGTGTCCCGACGG 771
DB |||||
QY 686 GGAGGAATATCAGCGGCTGTTTAAACGCGATTAATCTGTTGGTAGGCGTGTGAACAG 745
DB |||||
QY 772 AGACCAACATATCATCTGTGTCTGTTCTGTCTCAGTTCAGCGCGGTGGAATCGGCCCTGA 831
DB |||||
QY 746 AGACAAATATTCTCTTCTGTGTGGTTTCCGTTAGCTCAGCTGCTGTGNACTTGTCTCTGA 805
DB |||||
QY 832 TGAAGCTTCGGAAGTCTGAAGCACTGTCTAGCTAGGATGCTTCTGATGCTGTGCTGTAATA 891
DB |||||
QY 806 TGAAGCTTCGCGAGTCTTCAACAGTATCTGTCTAGGATGCTGTAAATGCGCTGGAAAAA 865
DB |||||
QY 892 TGGGAAAGCTAGTGTATCAACATCTGTTTGCACAAAGGATGCAAGAGGTTGTGTGGTGA 951
DB |||||
QY 866 TGGGAGAGCTTGTGATTAGCAATTTGATGGCTTAAGGTTGCAAAAAGTGTGTGGTAGTCA 925
DB |||||
QY 952 ACCGCTCCGTGGAAAGGGTGGATGCTATTTCGTGAGGAGATGAAGATATAGAGATCGTGT 1011
DB |||||
QY 926 ACAGAAAGTGAAGAGAGGGTTTTCAGCTATCCGTGAGGAGATGCTGTATTGAGATTATAT 985
DB |||||
QY 1012 ACAGCGCTCTTCAGACATGATCAAGCTGTCTGCTGTAAGCTGATGCTGTGTTTCAACAGCA 1071
DB |||||
QY 986 ATCSACCTCTTGATGAGATGCTAGCTTGTGCTTCAGAAAGCGGATGTTGTGTTTACCAGCA 1045
DB |||||
QY 1072 CGCATCTGAAACTTTCATTGTTTCGAAAGAACACGACAGAGCACTCCCGCTGTCTCTG 1131
DB |||||
QY 1046 CAGCCTCTGAGACGCCATTTGTTTAAAGGAGCATGTAGAGAAATCTCCCTCAAGCTTCTC 1105
DB |||||
QY 1132 ATACTATGGAGGTTTCCCTGTTTGTGCAATATCTGTCCCGAGGAATGTCAAGCGCAT 1191
DB |||||
QY 1106 CAGAGGTTGGAGGACTGAGGCATTTTGTGACATCTCTGTTCCGAGAAATGTTGATCTT 1165
DB |||||
QY 1192 GTGTCTCTGAAGTTGGCGTGCACGAGTGTACAATGTGACACACTTGAAGAGTGGTGG 1251
DB |||||
QY 1166 GTGTTGGTGAAGTGAAGAACAGCAGCGGGTTTACAATGTGGACCATCTCAAGAAAGTGTGG 1225
DB |||||
QY 1252 AAGCCAAACAGGAGGACCGGCTCAGGAAAGCAATGGAGCGGACAGCAATCATCACCGGAAG 1311
DB |||||
QY 1226 CTGCCAACAAAGACAGGATGAGGAAAGCAATGGAGCTCAACCATTAATCAAGAGG 1285
DB |||||
QY 1312 AACTGAGAGGTTTCAGGATGGAGGACTCGCTCGAGACCGTTCCGACCATCAAGAAAGC 1371
DB |||||
QY 1286 AATCCACACAGTTTGAAGCATGGAGGATTCATTGGAGACGGTTCACACGATCAAGAAAT 1345
DB |||||
QY 1372 TGAGTGTGTACGCGACAGGATCAGGGCTTCGAGCTCGAGAGTGGCTTCGAGAAAGTAG 1431
DB |||||
QY 1346 TAAAGAGCTTATGACAGAGAAATCAGAGTGGCAGAGCTCGAGAAAGTGCATGTCCAAAATGG 1405
DB |||||

QY 1432 GTGAGGACGCCCTCACCAAGAGATGAGGAGCCATCGAGGAGCTGAGCCGCGCATCG 1491
DB |||||
QY 1406 GAGATGA---TATCAACAAGAAAAACAACGAGAGCGGTTGATGACTTAAGCCGAGGTATCG 1462
DB |||||
QY 1492 TTAACAAGCTCTCTCATGGCCCGCTGACGACACTGAGGTGCGACGCGACGACGACCCGCA 1551
DB |||||
QY 1463 TGAACAGATTCTTACATGGTCCATGACGATTTGAGATGTGACCGNAGTGACAGTAGAA 1522
DB |||||
QY 1552 CCCTTGACGAGCGCTCGAGAACATGACGCGCTCAACCGGATGTTTCAGCTTCGACATGG 1611
DB |||||
QY 1523 CGCTGAGCGAGACCCCTTGAGAAATATGATGCTTTTGAACAGAAATGTACGCTC-----TTG 1576
DB |||||
QY 1612 AGRAGGCGATCATCGAGCAGAGATCAGGCGCCNAGGTGAGAGACACAAAACACTGA 1667
DB |||||
QY 1577 AGAAAGACATTTTGGAGGAAAAAGCTTAAGGCAATGCGCGAAACAACAACAGAAAGTAA 1632
DB |||||

RESULT 2

US-09-878-574-3208
; Sequence 3208, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 3208
; LENGTH: 384
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3028-014-Q1-B1-D9
US-09-878-574-3208

Query Match 6.5%; Score 125.6; DB 10; Length 384;

Best Local Similarity 64.6%; Pred. No. 4e-24;

Matches 204; Conservative 0; Mismatches 109; Indels 3; Gaps 1;

QY 1366 AGAGCTCAGGTGCTGACCGGACAGGATCAGGGCTCGAGAGCTCGAGAGTGCCTGCAG 1425
DB |||||
QY 1426 AAGTAGGTGAGGACGCCCTCACCAAGAGATGAGGAGCCATCGAGGAGCTGAGCACCG 1485
DB |||||
QY 61 AGATGGGTGATGATATAC---CAAAGAAAACCGGAGAGCTGTGATGACCTTAGTCGGG 117
DB |||||
QY 1486 GCATGTTAAACAGCTCTCCATGCGCCGCTGCAGCACTGAGGTGCGACGCGACGACGA 1545
DB |||||
QY 118 GTATAGTGAATAAGTTGCTTTCATGGTCCAATGCAACATTTAAGGTGTGATGGGAACGCA 177
DB |||||
QY 1546 GCCGACCCCTTGAACAGAGCTCGAGAACATGACGCGCTCAACCCGATGTTTCAGCTCG 1605
DB |||||
QY 178 GCCGACTCTTAGTGAGACACTGGAGAACATGAATGCTTTGAATAGGATGTTTCAACCTTG 237
DB |||||
QY 1606 ACATGGAGAACGCGATCATCGAGCAGAGATCAAGGCCCAAGTGGAGAGACACAAAAC 1665
DB |||||
QY 238 AGACNAAATATCTGTTTGGAGGAGAGATTCGAGCAAGAGTTCGAAACAAACCAAGAAAT 297
DB |||||
QY 1666 GAGGCCAGGAACCAAT 1681
DB |||||
QY 298 GAAATCTAACCAAT 313
DB |||||

RESULT 3

US-09-878-574-5465
; Sequence 5465, Application US/09878574
; Patent No. US20020110548A1

: GENERAL INFORMATION:		
: APPLICANT: Byrum, Joseph R.		
: APPLICANT: La Rosa, Thomas J.		
: APPLICANT: Thompson, Michael D.		
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with		
: TITLE OF INVENTION: Plants		
: FILE REFERENCE: 38-21(15401)B		
: CURRENT APPLICATION NUMBER: US/09/878,574		
: CURRENT FILING DATE: 2001-12-21		
: PRIOR APPLICATION NUMBER: 09/333,535		
: PRIOR FILING DATE: 1999-06-14		
: NUMBER OF SEQ ID NOS: 15775		
: SEQ ID NO 5465		
: LENGTH: 275		
: TYPE: DNA		
: ORGANISM: Glycine max		
: OTHER INFORMATION: Clone ID: 701097007H1		
US-09-878-574-5465		
Query Match	6.4%;	Score 122.6; DB 10; Length 275;
Best Local Similarity	71.4%;	Pred. No. 2.2e-23;
Matches 175;	Conservative 0;	Mismatches 69; Indels 1; Gaps 1;
QY	603	CATCTGTTTGAGGTGTGAGCTGGCTTGACCTTTTGGTTCTCGGTGAAGGACAAATCCTT 662
Db	7	CACTATTGTAAGTGGCTCAGGCGCTTGACCTACTTGTCTTGGGGAAGGTCAAATCTT 66
QY	663	GCTCAGGTTAAACAAGTTGTGAGGAGTGGACAGAACACAGTGGAGGCTTCGGAAAGAACATC 722
Db	67	GCTCAGGTGAAGCAGGTTGTGAAGCTGGACAGGAGTGCCTGGTTTGTATAGGAATC 126
QY	723	GATAGGATGTTCAAGGATGCAATCACTGCTGGAAAGCTGTCCGACGACGACCAACATA 782
Db	127	AGTGGTTGTTTCAAGCAGCGCATATCGTTGGGAAGCGGTTAGAACCGGACTAACATT 186
QY	783	TCATCTGGTCTGTTCTGTGTCAGTTCAGCGCGGTTGAACTGGCCCTGATGAAGCTTCG 842
Db	187	TCATCTGGATCAGTTCCTGTAAGCTCGCTGCTGTGGAGCT-GCACTGATGAAGCTACCG 245
QY	843	AAGTC 847
Db	246	GATTC 250
RESULT 4		
US-09-974-300-974		
: Sequence 974, Application US/09974300		
: Patent No. US20020146721A1		
: GENERAL INFORMATION:		
: APPLICANT: Berka, Randy M.		
: APPLICANT: Clausen, Ib Groth		
: TITLE OF INVENTION: Methods For Monitoring Multiple Gene		
: TITLE OF INVENTION: Expression		
: FILE REFERENCE: 10085-500-US		
: CURRENT APPLICATION NUMBER: US/09/974,300		
: CURRENT FILING DATE: 2001-10-05		
: PRIOR APPLICATION NUMBER: 09/680,598		
: PRIOR FILING DATE: 2000-10-06		
: PRIOR APPLICATION NUMBER: 60/279,526		
: PRIOR FILING DATE: 2001-03-27		
: NUMBER OF SEQ ID NOS: 8481		
: SOFTWARE: FastSEQ for Windows Version 4.0		
: SEQ ID NO 974		
: LENGTH: 1360		
: TYPE: DNA		
: ORGANISM: Bacillus licheniformis		
US-09-974-300-974		
Query Match	4.5%;	Score 86.6; DB 10; Length 1360;
Best Local Similarity	44.7%;	Pred. No. 5e-13;
Matches 485;	Conservative 0;	Mismatches 579; Indels 21; Gaps 3;
QY	315	ATAGCTGTGTATGGCCCTCAGTGTATACACAGACCACTGGAGATGCGTGAATAAATTGCT 374

```
Db      1066 CAGGC 1070

RESULT 5
US-09-974-300-5501
; Sequence 5501, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5501
; LENGTH: 1131
; TYPE: DNA
; ORGANISM: Bacillus clausii
US-09-974-300-5501

Query Match      4.3%; Score 83; DB 10; Length 1131;
Best Local Similarity 46.9%; Pred. No. 4.5e-12;
Matches 380; Conservative 0; Mismatches 410; Indels 21; Gaps 3;

QY      555 GAGCTTAGGAGCAGCTGTCATCTTCCGAGCAGTGATGCCACAGCCATCTGTTGAG 614
Db      47 GAGCTAACGCCGATTTGGTTATCCGTGAAATGACCATGCCATTAACACCTTTCCGC 106
QY      615 GTGTCAGCTGGCCTTGACTCTTTGGTTCTCGGTGAAGGACAAATCTTGTCAGGTTAA 674
Db      107 GTTGCTTGTGGTTAGACTATAGTGATCGCGGAAACGCAATTTCTGGACAGTGAAG 166
QY      675 CAAAGTTGTAGGAGTGGACGAACAGTCAGGAGGTTGGGAAGAACATCGATAGGATGTTTC 734
Db      167 CAGGCATTTTGTGTTGCCCAAGAGAGAAAGTCACAGGCACTGTGTTAAACAGCTATTT 226
QY      735 AAGGATGCAATCACTGCTGGAAGCGTTCGCGAGCGAGACCAACATATATCTGTGTCT 794
Db      227 AAACAAGCGGTTACGTCGCGCAACCGTGTGCACTCGTCAACGGAAATTTCTTCAACAGCC 286
QY      795 GTTCTGTGTCAGTTTCAGCGGCGGTTGAATCGGCCCTGATGAAGCTTCCGAACTCTGAAGCA 854
Db      287 GTTCTGTGTCATGCTATGCGCGGTAGAACTGGGAAGAAATTTTCGCGCAGTTTTCAGGA 346
QY      855 CTGTCAAGTAGGATGCTTCTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 914
Db      347 AAACATG-----TGCTTATTTTAGCGCTGGAAAAATGAGCGAACTGACGGCCAAACAT 400
QY      915 CTGGTTGCCAAGGATGCAAGAGGTTGTTGTGTCGAACCGCTCCGTGGAAAGGGTGGAT 974
Db      401 CTTTATTCGATGGGCGTCTGCTTCCATAGTGCATGAACCGGAAAGAAAGAAAGCGGTG 460
QY      975 GCTATTCTGTGAGGAGATGAAGATATAGAGATCGTGTACAGGCGCTCTCTCAGACATGTAT 1034
Db      461 GAACTGGCAAGCCAGTTTCTGGAACGGCT-----CGGTCTTTTGTCTGAATTGAAT 511
QY      1035 CAACTGCTGCTGAAGCTGATGTCGTTGTTACAGCAGCAGCGCATCTGAAACTTCATGTTGTC 1094
Db      512 GATGCTTTTAAAGAACGACATGTTGATAGTCTCCACGGGGCGCGGACCTTTGTTGTC 571
QY      1095 GCAAAAGAACACGACGAGGCACTCCCGCTGCTCTGTACTATGATGAGGAGTGTTCGCCGTG 1154
Db      572 AAAAGGAATATGCGCAGCTGCC-----TTAAAAACGGAAGAGCGAGCGCGTGT 625
QY      1155 TTTGTGACATATCTGTCCCGAGGAATGTGAGGCAATGTGTCGAAAGTGTGCGCTGCA 1214
```

```
Db      626 GCGATTGATATCGCAGTGCCACCGGATATTGACCCGGAATTTGCAACAAATTAGCGAGCTT 685
QY      1215 CGAGTTGTACAATGTTCGACGACTTTGAAAGAGGTGGTGAAGCCAAAGAGGACCGGCTC 1274
Db      686 TACTTATACGACATTTGATGACTTTACAGAACATTTGTTGAAGCCAAATAAAAAAGAGCGCAAA 745
QY      1275 AGAAAGCAATGAGGCGGCGAGACATCATCAGGAAAGAACTGAGACGGTTCGAGGCGATGG 1334
Db      746 AAAGAGGCTGAAAAAATCGGTGTTATGATTGAGGCGAGAGATCGATGACTTTTAAAGCATGG 805
QY      1335 AGGGACTCGCTCGAGACCGCTTCGACCATCA 1365
Db      806 CTAACACACTAGGCGGTGTGCTCTTATTA 836

RESULT 6
US-09-878-574-6807
; Sequence 6807, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 6807
; LENGTH: 271
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 701099031H1
US-09-878-574-6807

Query Match      3.5%; Score 68; DB 10; Length 271;
Best Local Similarity 62.2%; Pred. No. 2.7e-08;
Matches 107; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY      1494 ACAAGCTCTCCATGGCCCGCTGCAGCACCTTGAGTGCAGCGGACGACAGCCGACC 1553
Db      10 AATAAGCTACTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 69
QY      1554 CTTGACGAGACGCTCGAGAACATGCAAGCCCTCAACCGGATGTTTCAGCTTCGACATGGAG 1613
Db      70 CTGAGTGAAGTACTTTGAGATATGCGGCCCTTAACAGATGTTATGATCTTGACAGAA 129
QY      1614 AAGGCGATCATCGACAGAAAGATCAAGGCCAAAGGTGGAGAAAGACAAAACT 1665
Db      130 ACTTCTTGATCGAAGAAAGATCAGAGTCAAGATGGAACGGGTTTCAAGAT 181

RESULT 7
US-10-156-761-4723
; Sequence 4723, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
```


;; PRIOR APPLICATION NUMBER: JP 2001-272697
;; PRIOR FILING DATE: 2001-08-02
;; NUMBER OF SEQ ID NOS: 15109
;; SEQ ID NO 4723
;; LENGTH: 1707
;; TYPE: DNA
;; ORGANISM: Streptomyces avermitilis
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (1)..(1707)
US-10-156-761-4723

Query Match
Best Local Similarity 3.1%; Score 60.6; DB 15; Length 1707;
Matches 168; Conservative 0; Mismatches 179; Indels 0; Gaps 0;
QY 435 GAGGCTGCTTCTTAGTACCTGTAATAGAAATTTATGTGTGGCGCTATCATGG 494
DB 127 GAGGCGCGGCTCTTCCACCTGCAACCGCATCGAGCTGTACGCCAGCTGACAAATTC 186
QY 495 AACCGTGGTATCAGAGAACTAGTGGACTGTGATGTCGAAGAAAGTGGTATTCGCCGCTTCC 554
DB 187 CACGCGGCGTGGCGAGCTGTCCAGCTGTCTGCCACACACAGCGGGGTCTGGGCTCGAG 246
QY 555 GAGCTTAGGAGACCTCTTTCATCTTTGCGAAGCAGTGTGCCACACGCCCATCTGTTTGA 614
DB 247 GAGCTCACTCCCTATCTCTACGTGCACTACGAGGACCGGGCGCTCCACCACTCTTCTCG 306
QY 615 GTGTACGTGGCTTGGACTCTTTGTTCTCGGTGAAGCAAAATCTTCTCAGGTTAA 674
DB 307 GTGGCTGGGCTCGACTGTGATGTTGTCGTCGAGGCGGAGCGCTCTCAACGACTGTT 426
QY 675 CAAGTTGTGAGGAGTGGACAGACAGTGGAGGCTTGGGAAAGAACTGATAGGATGTT 734
DB 367 GACGCGTGGCCACCGCGAGAACTGACACCGCGGAGCGCTCTCAACGACTGTT 426
QY 735 AAGGATGCAATCACTGTCTGGAAGGCTGTCGCGAGGACCAACAT 781
DB 427 CAGCAGGCACTGGGACCGGCAAGCGGCGCACTCCGAGACCGGCAT 473

RESULT 8

US-10-156-761-1
;; Sequence 1, Application US/10156761
;; Publication No. US20030119018A1
;; GENERAL INFORMATION:
;; APPLICANT: OMURA, SATOSHI
;; APPLICANT: IKEDA, HARUO
;; APPLICANT: ISHIKAWA, JUN
;; APPLICANT: HORIKAWA, HIROSHI
;; APPLICANT: SHIBA, TADAYOSHI
;; APPLICANT: SAKAKI, YOSHIYUKI
;; APPLICANT: HATTORI, MASAHIRA
;; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;; FILE REFERENCE: 249-262
;; CURRENT APPLICATION NUMBER: US/10/156,761
;; CURRENT FILING DATE: 2002-05-29
;; PRIOR APPLICATION NUMBER: JP 2001-204089
;; PRIOR FILING DATE: 2001-05-30
;; PRIOR APPLICATION NUMBER: JP 2001-272697
;; PRIOR FILING DATE: 2001-08-02
;; NUMBER OF SEQ ID NOS: 15109
;; SEQ ID NO 1
;; LENGTH: 9025608
;; TYPE: DNA
;; ORGANISM: Streptomyces avermitilis
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (4187715)
;; OTHER INFORMATION: a, t, c, g, other or unknown

Query Match
Best Local Similarity 3.1%; Score 60.6; DB 15; Length 9025608;

Best Local Similarity 48.4%; Pred. No. 0.0014;
Matches 168; Conservative 0; Mismatches 179; Indels 0; Gaps 0;
QY 435 GAGGCTGCTTCTTAGTACCTGTAATAGAAATTTATGTGTGGCGCTATCATGG 494
DB 5777935 GAGGCGCGGCTCTTCCACCTGCAACCGCATCGAGCTGTACGCCAGCTGACAAATTC 5777994
QY 495 AACCGTGGTATCAGAGAACTAGTGGACTGTGATGTCGAAGAAAGTGGTATTCGCCGCTTCC 554
DB 5777995 CACGCGGCGTGGCGAGCTGTCCAGCTGTCTGCCACACACAGCGGGGTCTGGGCTCGAG 5778054
QY 555 GAGCTTAGGAGACCTCTTTCATCTTTGCGAAGCAGTGTGCCACACGCCCATCTGTTTGA 614
DB 5778055 GAGCTCACTCCCTATCTCTACGTGCACTACGAGGACCGGCGCTCCACCACTCTTCTCG 5778114
QY 615 GTGTACGTGGCTTGGACTCTTTGTTCTCGGTGAAGCAAAATCTTCTCAGGTTAA 674
DB 5778115 GTGGCTGGGCTCGACTGTGATGTTGTCGTCGAGGCGGAGCGCTCTCAACGACTGTT 5778174
QY 675 CAAGTTGTGAGGAGTGGACAGACAGTGGAGGCTTGGGAAAGAACTGATAGGATGTT 734
DB 5778175 GACGCGTGGCCACCGCGAGAACTGACACCGCGGAGCGCTCTCAACGACTGTT 5778234
QY 735 AAGGATGCAATCACTGTCTGGAAGGCTGTCGCGAGGACCAACAT 781
DB 5778235 CAGCAGGCACTGGGACCGGCAAGCGGCGCACTCCGAGACCGGCAT 5778281

RESULT 9

US-08-781-986A-274
;; Sequence 274, Application US/08781986A
;; Publication No. US20030054436A1
;; GENERAL INFORMATION:
;; APPLICANT: Charles Kunsch
;; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
;; NUMBER OF SEQUENCES: 5255
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Human Genome Sciences, Inc.
;; STREET: 9410 Key West Avenue
;; CITY: Rockville
;; STATE: Maryland
;; COUNTRY: USA
;; ZIP: 20850
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;; COMPUTER: HP Vectra 486/33
;; OPERATING SYSTEM: MSDOS version 6.2
;; SOFTWARE: ASCII Text
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/781.986A
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Benson, Bob
;; REGISTRATION NUMBER: 30,446
;; REFERENCE/DOCKET NUMBER: PB248PP
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (301) 309-8504
;; TELEFAX: (301) 309-8512
;; INFORMATION FOR SEQ ID NO: 274:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1496 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
US-08-781-986A-274

Query Match
Best Local Similarity 49.2%; Pred. No. 2e-05;
Matches 185; Conservative 1; Mismatches 184; Indels 6; Gaps 1;

```
QY 591 GATGCCACAGCCATCTGTTGAGGTGTCAGCTGGCCTTGACTCTTTGGTTCTCGGTGAA 650
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
777 GAAGCAGTAGAACATTTATTGCGTGTCACTTCTGGTTTAGATTCAATCGTACTTTGGAGAA 836
QY 651 GGACAAATCCTTGCTCAGGTAAACAAAGTTGTGAGGAGTGGACAGAAACAGTGGAGCTTG 710
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
837 ACTCAAAATTTAGTCAAAATAGAGATGCAATTTCTTAGCCGCAAGACAGGTAGACA 896
QY 711 GGAAGAACATCGATAGGATGTTCAAGGATGCAATCACTGCTGGAAAGCTGTCGCGACG 770
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
897 GGRACAAATTTTAAATCATCTATTAAACAGGCAATTTACTTTTGCAAAAAGACATAAT 956
QY 771 GAGACCAACATATCATCTGCTGCTGTTTCTGTCAGTTCAGCGCGGTTGAACTGGCCCTG 830
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
957 GAAACAGATATAGCTGATTAATCTGTAAGTGTGCTTATGCTGCGGTGAGTTGGCGAAA 1016
QY 831 ATGAAGCTTCCGAAGTCTGAAGCAGTGTGAGTGGATGCTTCTGATTGGTCTCTGTAAA 890
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1017 AAAGTATTGGCAAAATTTGAAAAGTAGCAAGCTA-----TCATTATTGGTGCAGGGAA 1070
QY 891 ATGGGAAAGCTAGTATCAAAATCTGTTGTCGCAAGGATGCAAGAGTTGTTGTGGTG 950
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1071 ATGAGTGAATTAATCACTATTAAATCTTCTTGGTCTGGAATTAATGATATTACAGTAGTA 1130
QY 951 AACCGCTCCGTGGAAA 966
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1131 AATAGAACAAATTTGAAA 1146
```

RESULT 10

```
US-09-770-791-195
; Sequence 195, Application US/09770791
; Patent No. US20020062014A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woesner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Krickler, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: thaliana
; CURRENT APPLICATION NUMBER: US/09/770,791
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,480
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 195
; LENGTH: 380
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-791-195
```

Query Match 2.9%; Score 56.2; DB 9; Length 380;
Best Local Similarity 69.2%; Pred. No. 5.9e-05;
Matches 92; Conservative 0; Mismatches 38; Indels 3; Gaps 1;

```
QY 235 CGTGGCAAGCTGCCAGCTCGCCGCTCGAGCAGTTCAAGATAT---CCGCGGACC 291
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
247 CTGCGTCTAAATGACAGTAGCATCTCTGCTCTTGAAACAACTCAAGAACTCTGCGAGCTGATC 306
QY 292 GGTACATCAAGAAAGGAGTAGTACCATAGCTGTGATAGGCTCAGTGTACACACAGACACAG 351
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
307 GATATACAAGAAAGAACAGCAGTATTGTTGTGATTGGACTTAGTATTACACAGCTCCTG 366
QY 352 TGGAGATCGGTGA 364
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
367 TTGAGATCGGTGA 379
```

RESULT 11

```
US-09-738-626-470
; Sequence 470, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 470
; LENGTH: 1389
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-470
```

```
Query Match 2.8%; Score 54.4; DB 10; Length 1389;
Best Local Similarity 44.4%; Pred. No. 0.0004;
Matches 220; Conservative 0; Mismatches 276; Indels 0; Gaps 0;

QY 343 CAGCACCAGTGGAGATGGTCAAAAACTTGCTGTGTCAGAGAGAACTGTGGCCCCGTGCTA 402
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
35 CTGCGCTGTGTCGCTCTTCTTGACGCTGAGCATGGATGATTCAGTACGTGGTGAACAA 94
QY 403 TTCAGAACTCAGTACAGCTGAACCATATTGAAGAGGCTGCTGTTCTTAGTACCTGTAATA 462
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
95 CTCAGACACTCTGGTAGGGCGTCTTTAAGCGAGGCCCTCATTTGTTCTCTAGCTGTAACC 154
QY 463 GAATGGAAATTTATGTGTGGCGCTATCATNGAAACCGTGGTATCAGAGAAGTAGTGGACT 522
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
155 GCCTGGAGGTCTACACCGTCACTAGCAGTTTCCATATCGTGTGTTAATGATGTGGTGGAGG 214
QY 523 GGATGTCGAGAAAAGTGGTATTCGCCGCTCCGAGCTTAGGGAGGACCTGTTTCATCTTGC 582
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
215 TTCTCCATGAGGCAAGTGGCGGTAGATATTGAAATCTTTCGCGGGATATCTTTTATGTCGGT 274
QY 583 GAAGCAGTGTGCCACACGCCATCTGTTTGAAGGTGTGAGCTGGCGCTTGACTCTTTGGTTC 642
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
275 ACGCGGATGCTGCTGTGACACACATGTTGTGGTGTGACTTCCGGTGTGGATTCCATGGTGT 334
QY 643 TCGGTGAAGGCAAAATCTTCTGCTCAGGTATAACAAAGTTGTGAGGAGTGGACAGAACAGTG 702
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
335 TGGGTGAGCAGCAGATCATTTGGTCAGGTGGCGCACTGCGTACCAAGCAGCTAATGAATATG 394
```

QY 703 GAGGCTTGGAAAGAACATCGATAGATGTTCAAGATGCAATCACTCTCGAAGCGTG 762
 Db 395 GTTCTGTCGCTCTGCTTTCATTCATTCACCAACCGCGTGCATACCGCAAGCGG 454
 QY 763 TCCGAGGAGAGAACCAATATCATCTGCTGCTGTTTCTGTCAGTTTCAGCGCGGTGAAC 822
 Db 455 TGCATTCCGAGACTCTGATTGATGCTGCTGTCATCATGCTGCTGCTTTCGCTGTGGATC 514
 QY 823 TGGCCCTGATGAGCT 838
 Db 515 GCGCGTTGGTCAGAT 530

RESULT 12
 US-09-738-626-1
 ; Sequence 1, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAOKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738, 626
 ; CURRENT FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: PatentIn ver. 3.0
 ; SEQ ID NO 1
 ; LENGTH: 3309400
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 US-09-738-626-1

Query Match 2.8%; Score 54.4; DB 10; Length 3309400;
 Best Local Similarity 44.4%; Pred. No. 0.041;
 Matches 220; Conservative 0; Mismatches 276; Indels 0; Gaps 0;

QY 343 CAGCACCAGTGGAGATGGTGAAGAACTTGTCTGTCAGAGGAATGTGGCCCGTGTCTA 402
 Db 438550 CTGCGCTGTCTGCTCTTCTGAACGCTCTGAGCATGGATGATTCAGTACGTGGTGAACAA 438609
 QY 403 TTCAAGAACTCACTAGCTGAACCATATTTGAAGAGGCTGCTGTTCTAGTACCTGTAATA 462
 Db 438610 CTCAGGACTCTCTGGTAGGCGCTTTTAAAGGAGGCGCTCATTTCTCTACGTGTAAAC 438669
 QY 463 GAATGGAAATTTATGTGTGGCGCTATCATGAAACCGTGTGTATCAGAGAACTAGTGGACT 522
 Db 438670 GCCTGGAGGTCTACACCGTCACTAGCAGTTTCCATACCTGTTGTTAATGATGTGTGGAGG 438729
 QY 523 GGATGTCGAGAAAGTGTATTCCTCGCTTCCGAGCTTAGGAGCACTGTTTCACTTTTC 582
 Db 438730 TTCTCCATGAGGCAAGTGGCGTAGATATGAAACTTTTCCGCGGATATCTTTATGTCGCTT 438789
 QY 583 GAAGCAGTGTATGCCACACCGCATCTGTTTGAAGTGTGTCAGCTGGCTTGACTCTTTGGTTC 642
 Db 438790 AGCCGATGCTGCTGCTGAACACATGTTGGTGGTACTTCGGGTGGATTTCCATGTTGT 438849
 QY 643 TCGGTGAAGACAAATCTTGTCTCAGGTAAACAAAGTTGTGAGGAGTGGACAGAACAGTG 702

Db 438850 TGGGTGACGACGAGATCATTGGTCAGGTGCGCACTGCGTACCAAGCAGCTAATGAATATG 438909
 QY 703 GAGGCTTGGAAAGAACATCATGATAGGATGTTCAAGATGCAATCACTCTGTAAGAGCGTG 762
 Db 438910 GTTCTGTCGCTCTGCTTTCATTCATTCACCAACCGCGTGCATACCGCAAGCGG 438969
 QY 763 TCCGAGGAGAGAACCAATATCATCTGCTGCTGTTTCTGTCAGTTTCAGCGCGGTGAAC 822
 Db 438970 TGCATTCCGAGACTCTGATTGATGCTGCTGTCATCATGCTGCTGCTTTCGCTGTGGATC 439029
 QY 823 TGGCCCTGATGAGCT 838
 Db 439030 GCGCGTTGGTCAGAT 439045

RESULT 13
 US-10-156-761-5384
 ; Sequence 5384, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: HORIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156, 761
 ; CURRENT FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 5384
 ; LENGTH: 3303
 ; TYPE: DNA
 ; ORGANISM: Streptomyces avermitilis
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(3303)
 US-10-156-761-5384

Query Match 2.7%; Score 51; DB 15; Length 3303;
 Best Local Similarity 50.6%; Pred. No. 0.0058;
 Matches 123; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 72 ATGCGGACGACGAGCTCAGCGACACCCGCGCGCGCAGCGCCACCCAGCCGCGGCGGCTCCGG 131
 Db 433 ATCGCGCTCGACGACCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 492
 QY 132 CGGGGTCGTCGTCGCGCGCTCTCCAGAGGGTGGCGCGCGCGCGCGCGCGCGCGCGGCTCCGG 191
 Db 493 ACGGCGACGCTACTCCGCGCGCTCCAGAGCACCCCGCGCGGAGAACCCGCTACACCCCGTCC 552
 QY 192 GTGTGTGGTGGCGGCG 251
 Db 553 TTGAGAGCGGCGCGCGCTCGACTTTCGTGAAGGACACCCCGGAGGCGGCGCGCGCGCGGTC 612
 QY 252 AGGTCGCCCGCTTCGAGCAGTTCAAGATATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 311
 Db 613 ACCATCGCGCTCTCGACTCGGCGCTGGAACCTCGGCGCGCGCGCGCGCGCGCGCGCGCG 672
 QY 312 ACC 314
 Db 673 ACC 675

RESULT 14
 US-10-156-761-1/c
 ; Sequence 1, Application US/10156761

Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 1
LENGTH: 9025608
TYPE: DNA
ORGANISM: Streptomyces avermitilis
FEATURE:
NAME/KEY: misc feature
LOCATION: (4187715)
OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match 2.7%; Score 51; DB 15; Length 9025608;
Best Local Similarity 50.6%; Pred. No. 0.64;
Matches 123; Conservative 0; Mismatches 120; Indels 0; Gaps 0;
QY 72 ATGGGACACACGACGCTCAGGACACCCGCGCGGAGCGCCACACCCGCGCGG 131
Db 6549578 ATCGGCTCGAGACCCGCGCGGAGCGCCGACCGCAAGGCGCGTCTCGAAGGCC 6549519
QY 132 CGGGGGTCTGCTGGCCCTCTGCCAGAGGCTGGCGCGGCGGAGCGCGGCTCCGGG 191
Db 6549518 ACGGGACGCTACTCCGCGCGCTCCAGAGCACCCCGCGGAGAACCCGTACACCGTCC 6549459
QY 192 GTGGTGGGTGGAGCGCCCGCGGCTGGAGGCGCGGAGCGCGGCGGCGGCGGCGG 251
Db 6549458 TTGAGAGCGGCGCGGCTGACTTCGTGAAGGACACCCGAGGCGGAGCGCGGCGGCG 6549399
QY 252 AGCGTCCGCGCGCTCGAGCAGTTCAAGATATCCGCGGCGGCTGACATGAAGGAGGAGT 311
Db 6549398 ACCATCGGCGTCTCGACTCGGCGTGGACCTCGGCGCGGCTCGAGAGCGGAC 6549339
QY 312 ACC 314
Db 6549338 ACC 6549336

RESULT 15
US-10-190-435-50
Sequence 50, Application US/10190435
Publication No. US20030143248A1
GENERAL INFORMATION:
APPLICANT: ZUR MEDEDE, Jan
APPLICANT: BARNETT, Susan W.
APPLICANT: LIAN, Ying
APPLICANT: ENGELBRECHT, Susan
APPLICANT: VAN RENSBURG, Estrelita J.
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
FILE REFERENCE: PP18133.003 / 2302-18133
CURRENT APPLICATION NUMBER: US/10/190,435
CURRENT FILING DATE: 2002-12-30
NUMBER OF SEQ ID NOS: 319
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 50
LENGTH: 1965
TYPE: DNA

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: protInart.YMM.opt_C
US-10-190-435-50

Query Match 2.6%; Score 50.6; DB 13; Length 1965;
Best Local Similarity 47.4%; Pred. No. 0.0055;
Matches 152; Conservative 0; Mismatches 169; Indels 0; Gaps 0;
QY 1334 GAGGACTCGCTGGAGACCGTTCCGACCATCAAGAGCTGAGGTCTGTACGCGGACAGGAT 1393
Db 1560 GTGGTACCAGCTGGAGAGGAGGCCCATCATCGGCGCCGAGACCTTTCTAGCTGACGGCGC 1619
QY 1394 CAGGSCCTCGGAGCTCGGAGAGTGCCTCGAGAAAGTAGTAGGAGCGCCTCACCAGAA 1453
Db 1620 CGCCAAACCGGAGACCAAGATCGGCAAGGCCGCTACGTGACCCGCGGCGCGGCGGCGG 1679
QY 1454 GATGAGGAGAGCCATCGGAGGAGCTGAGCACCGGCATCGTTAACAGCTCTCCATGGCCC 1513
Db 1680 GATCGTGGAGCTGACCGGAGACCCACACAGAGACCGAGCTGCGAGGCCATCCAGTGGC 1739
QY 1514 GCTGAGACCTGAGTGGAGTGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 1573
Db 1740 CCGGAGGAGCGGAGCGGAGGTTGAACATCGTGACCGAGCGGAGTACGCGCTGGCAT 1799
QY 1574 CATGACGCGCTCAACCGGATGTTTCCAGCTCGACATGAGAGAGGCGGATCATCGAGCAGAA 1633
Db 1800 CATCCAGGCGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1859
QY 1634 GATCAAGGCCAAGGTGGAGAA 1654
Db 1860 GATCAAGAGGAGAGGTGTA 1880

Search completed: December 15, 2003, 18:32:45
Job time : 1550 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 12:46:10 ; Search time 111 Seconds
(without alignments)
7650.641 Million cell updates/sec

Title: US-10-018-902-3

Perfect score: 1924

Sequence: 1 ccacgcgtccgcacatcaataa.....aaaaaaaaaaaaaaaaaag 1924

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*

2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*

3: /cgn2_6/prodata/2/ina/6A_COMB.seq:*

4: /cgn2_6/prodata/2/ina/6B_COMB.seq:*

5: /cgn2_6/prodata/2/ina/PCUTUS_COMB.seq:*

6: /cgn2_6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	87.4	4.5	1308	4	US-09-252-991A-13991
2	87.4	4.5	1338	4	US-09-252-991A-14144
3	82.6	4.3	1026	4	US-09-252-991A-14075
4	64.6	3.4	1347	4	US-09-134-001C-218
5	64.2	3.3	1664976	4	US-08-916-421B-1
6	57.4	3.0	1290	4	US-09-328-352-517
7	56.4	2.9	7218	1	US-08-232-463-14
8	53	2.8	4403765	3	US-09-103-840A-2
9	53	2.8	4411529	3	US-09-103-840A-1
10	50	2.6	8147	4	US-09-514-247A-9
11	49.6	2.6	42988	4	US-08-311-731A-128
12	49.4	2.6	1203	3	US-09-086-010-1
13	48.8	2.5	1149	2	US-08-466-103A-3
14	48.4	2.5	77536	4	US-09-410-551B-1
15	47.6	2.5	4403765	3	US-09-103-840A-2
16	47.6	2.5	4411529	3	US-09-103-840A-1
17	45	2.3	768	4	US-09-252-991A-12559
18	45	2.3	1377	3	US-09-377-557-17
19	45	2.3	1425	4	US-09-252-991A-12926
20	45	2.3	2002	4	US-09-016-434-1172
21	44.6	2.3	984	4	US-09-252-991A-14047
22	44.6	2.3	1488	4	US-09-252-991A-8426
23	44.6	2.3	2049	4	US-09-252-991A-8348
24	44.6	2.3	2118	4	US-09-252-991A-14221
25	44.6	2.3	2121	4	US-09-252-991A-14004
26	44.6	2.3	2205	4	US-09-252-991A-14132
27	44.4	2.3	5118	4	US-08-669-785-3

C 28	44.2	2.3	1935	4	US-09-620-312D-236	Sequence 236, App
C 29	44.2	2.3	2825	4	US-09-196-390-5	Sequence 5, Appli
C 30	43.8	2.3	1134	4	US-09-622-439-21	Sequence 21, Appli
C 31	43.8	2.3	1512	3	US-08-476-509B-1	Sequence 1, Appli
C 32	43.8	2.3	1626	3	US-08-348-518C-1	Sequence 1, Appli
C 33	43.6	2.3	744	4	US-09-368-819A-1	Sequence 3, Appli
C 34	43.6	2.3	795	4	US-09-368-819A-3	Sequence 3, Appli
C 35	43.6	2.3	1173	3	US-08-993-380-3	Sequence 3, Appli
C 36	43.4	2.3	1621	1	US-08-722-001-13	Sequence 13, Appli
C 37	43.4	2.3	4079	4	US-09-016-434-1237	Sequence 1237, Ap
C 38	43.2	2.2	2175	4	US-09-252-991A-15144	Sequence 15144, A
C 39	43.2	2.2	2588	2	US-08-796-414B-6	Sequence 6, Appli
C 40	43.2	2.2	2721	4	US-09-252-991A-14887	Sequence 14887, A
C 41	43.2	2.2	2781	4	US-09-252-991A-14482	Sequence 14482, A
C 42	43	2.2	494	4	US-09-056-556-176	Sequence 176, App
C 43	43	2.2	494	4	US-09-072-596-171	Sequence 171, App
C 44	43	2.2	1647	1	US-08-083-948-13	Sequence 13, Appli
C 45	43	2.2	1647	1	US-08-393-785-13	Sequence 13, Appli

ALIGNMENTS

RESULT 1

US-09-252-991A-13991

; Sequence 13991, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 13991

; LENGTH: 1308

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-13991

Query Match 4.5%; Score 87.4; DB 4; Length 1308;

Best Local Similarity 46.5%; Pred. No. 2e-10;

Matches 585; Conservative 0; Mismatches 636; Indels 36; Gaps 8;

QY	313	CCATAGCTGTGTAGCCCTCAGTGTACACAGCAGCAGTGGAGATGCGTGAACAACTTG	372
DB	44	CCTTCATTGCGCTCGGCATCAATCAAGACCGCCTCGGTGGCTGTCGCGAGCGGTGG	103
QY	373	CTGTTCAGAGAACTGTGGCCCGCTGCTATTCAAGAACTCACTAGCCCTGAACCATATTG	432
DB	104	CTTTCACCTCCGAAAGAGATGTCGAGCGCTGTCAGAGCTGTGCCCTCCACACAGCC	163
QY	433	AAGAGCTGCTGTTCTTAGTACCTGTAATAGAAATTTATGTGTGGCGCTATCAT	492
DB	164	GCGAGGCGGCCATCTGTGACCTGCAACCGAGCGAACTCTACCTGGAGAT-CGACCAT	222
QY	493	GGAACCGTGGTATCAGAGAGTAGTGGACTGATGCGAAGAAAGTGGTATTCGCGCTT	552
DB	223	CCGACCGCGGAC-----GACGTGCTCGCTGCTGCGGACTACCATCGCTGACTCTCG	277
QY	553	CCGAGCTTAGGAGCACCTGTTTCATCTTCGAAAGAGTATGCCACACGCCATCTGTTTG	612
DB	278	ACGATTTGGCGCCCTGTGCTTATGTGCACAGGAGAGACGCGGTGCCACATATGTC	337
QY	613	AGGTGTCAGCTGGCCCTTGAATCTTTGGTTCTCGTGAAGGACAAATCTTGTTCAGGTTA	672
DB	338	GCGTGGCTCGGCTCGACTCGATGATGTTGCTCGCGAGCGCGAGATCTCGCCAGATCA	397

673 AACAGTTGTGAGGAGTGGACAGACAGTGGAGGCTTGGAAAGAACATCATAGGATGT 732
Db AGTCGGCTACGGCGTGGCGGCGAGCGCCGACCGTGGTGGCTGGCTGGCGCGCTGT 457
733 TCAAGGATCAATCACTGTGGAAGCGTGTCCGAGCGGACGACCAACATATCATCTGGTG 792
Db TCCAGGCCACTTTCAGTACGCGAAGACCGCTCCGACCGATACCGCATCGCGAGAAC 517
793 CTGTTTCTGTCACTTCAAGCGCGGTGAACCTGGCCCTGATGAAGCTTCCGAATCTGAAG 852
Db CGGTGTCGGTGGCGGTGGCGCGGTGAGCTGCGCCCAAGCAG-ATCTTCAGCGACCTGCAC 576
853 CACTGTCACTAGGATGCTTCTGATTGGTGTCTGTAATAATGGAAAGCTAGTATCAAAAC 912
Db CGCAGCCAGC-----GCTGCTGATCGCGCCCGGAGACCATCACTGCTGGCGGCC 631
913 ATCTGTTGCCAAGAGATGCAAGAGTTGTTGTGTGAACCGCTCCGTGGAAAGGTTGG 972
Db ACCTGTTTCGAGCAAGCGTGAAGCGCATCGTGTGGCCAAACCGCACCTGGAGCGGCCA 691
973 ATGCTATTCTGTAGGAGATGAAGATATAGAGATCGTGTACAGGCTCTCTCAGACATGT 1032
Db GCTGCTCGCGAGCAGTTCGGCGCCCATCGGTGCTG-----CTTTCGAGATAC 742
1033 ATCAAGCTGCTGTAAGCTGATGTGCTGTTCACAGCACCGCATCTGAACCTTCATTGT 1092
Db CCGAGAGTTGGCCACAGCGCATATGTCATAGCTCCATGTCAGCGCATGTCGCCATCC 802
1093 TCGCAAAAGAACACGAGAGGCACTCCCGCTGTCTGTATATATATGGAAGTGTTCGCC 1152
Db TCGCAAGGCGCGCTGGAG-CGCGCGCTGAAGCAGCGCAAGCAACGCGATGTTCA-- 859
1153 TGTGTTCTGACATATCTGTCCCGAGGATGTGAGCATGTGTCTGAAGTTGGCGCTG 1212
Db ---TGTTCGACATCGCGCTGCCCGGACATCGAGCGGAACTGGCGAACTGGACGAG 916
1213 CACGAGTGTACAAATTCGACGACTTGAAGAGGTGTGGAAGCAACAAAGGAGGACCGGC 1272
Db TCTACCTCTATACGCTCGAGACCTCCACGAAGTGTGGCGGAGATCTCAAGACGCGC 976
1273 TCAGGAAAGCAATGGAGGCGCAGACATATCAACGAAGACTGAGACGTTTCGAGGCAT 1332
Db AGGCGCTGCCAGGCGCGGAGAACTGCTGGCAGCGCGCTGCGCGAGTTCATGACG 1036
1333 GAGGAGCTCGCTGGAGACCGTTCCGACCATCAAGAGCTGAGTACGCGGACAGGA 1392
Db GCTTCGCGMACTTGGCGCGTGGAGCTGCTGCGCGCTATCGGCGAGCGCGCGCC 1096
1393 TCAGGCGCTCGAGCTCGAAGTGTCTCGAAGAGTGTGAGGAGCGCCCTCACCAAGA 1452
Db TGGCGATGAGGAGCTGGGCAAGCGCCAGCGCACTGGCCACGCTGCGACCCCGCG 1156
1453 AGATGAGGAGAGCATCGAGGAGCTGAGCACCGGCATCGTTAAACAAGCTCTCCATGGCC 1512
Db AGGTG-----ATGGCCCACTGGCGCGCGCTGACCAACAAACTGCTCAGCGCG 1207
1513 CGCTGACGACCTGAGTGGAGCGGAGCGGAGCGGACCGCACCTTTCGAGGAGCTCG 1569
Db CCAGCGTGCAGATGAAGAAATGTCGCGAGGCGCGCATCGAGCGCTGGCGCTCG 1264

RESULT 2

US-09-252-991A-14144/C
; Sequence 14144, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14144
; LENGTH: 1338
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14144

Query Match 4.5%; Score 87.4; DB 4; Length 1338;
Best Local Similarity 46.5%; Pred. No. 2.1e-10;
Matches 585; Conservative 0; Mismatches 636; Indels 36; Gaps 8;

QY 313 CCATAGCTGTATAGGCTCAGTGTACACACAGCACAGTGGAGATCGGTGAAAAATTG 372
Db CTTTCAATGCGCTCGGCATCAATCAACAGACCGCTCGGTGGTGTCCGAGCGCGTGG 1257
QY 373 CTGTTGCAGAGGAACGTGGCCCGTCTATTCAAGAACTCCTAGCCTGAACCATATTG 432
Db CTTTCACTCCGGAACAGATGGTCAAGCGCTGCGAGCTCTGCCGCTGACCACCGACC 1197
QY 433 AAGAGCTGCTGTTCTTAGTACCTGTAATAGAAATGAAATTTATGTGGTGGCGTATCAT 492
Db GCGAGCGGCGCATCTGTGACCTGCAACCGAGCGAATCTTACCTGGAGAT-CGACCAT 1138
QY 493 GGAACCGTGTATCAGAGAGTAGTGAATGTGGAAGAAAGTGTATTTCGCGCTT 552
Db CCGACCGCGGAC-----GACGTGCTCGCTGGCTGGGACTACCATCGCTGACTCTCG 1083
QY 553 CCGAGCTTAGGAGACCTGTTTCATCTTGGAGAGCAGTATGCCACACCGCATCTGTTG 612
Db ACAGTGTGCGCGCTGTGCTATGTGCAACGAGACGAGCGGTCGCGCATGATGC 1023
QY 613 AGTGTGAGTGGCTTGAATCTTTCGTTCTCGTCAAGGACAAATCTTGTCTCAGGTTA 672
Db GCGTGGCTCGCGCTCGACTCGATGCTCTCGCGAGCGGAGATCTCGGCGCAGATGA 963
QY 673 AACAGTTGTGAGAGTGAAGAGAAACAGTGGAGGCTTGGAAAGAAACATCATGATGATG 732
Db AGTGGCGCTACGCGCTCGCGCGAGCGCGACCGTGGTGGCTGCTGGCGCGCTGT 903
QY 733 TCAAGGATCAATCACTGCTGGAAGCGTGTCCGAGCGGAGACCAACATCATCTGTTG 792
Db TCAGGCGCACTTCAGTACGCGGAAGACCGTTCGCGACCGATACCGCATCGGCGAGAAC 843
QY 793 CTGTTTCTGCTCAGGCTTCAAGCGCGGTTGAATCTGCTGATGAAGCTTCCGAAAGTCTGAAG 852
Db CGGTGTCGCTGGCGTTCGCGCGGTGAGCTTGGCCAAAGCAG-ATCTTCAGCGACCTGCAC 784
QY 853 CACTGTCACTAGGATGCTTCTGATTTGGTGTGGTAAATGGGAAAGCTAGTATCAAAAC 912
Db CGCAGCGAGG-----GCTGCTGATCGCGCGCGCGGAGACCATCACCTGCTGGCGCGCC 729
QY 913 ATCTGTTGCCAAGAGATGCAAGAGTGTGTTGTGTGAACCGCTCCGTGGAAAGGTTGG 972
Db ACTGTTTCGAGCAAGCGGTGAAGCGCATCGTGTGGCGCAACCGCACCTTGGAGCGCGCCA 669
QY 973 ATGCTATTCTGTGAGGAGATGAAGATATAGAGATCGTGTACAGGCGCTCTCTCAGACATGT 1032
Db GCTGCTCGCGGAGCAGTTCGGCGCCCATCGGTGCTG-----CTTTCGAGATAC 618
QY 1033 ATCAAGCTGCTGTAAGCTGATGTGTTTCAAGGACCGCATCTGAACATTCATTTGT 1092
Db CCGAGGAGTTGGCCAAACGCGATATCGTTCATCAGCTCCACTGCGAGCGAGTTCGCCATCC 558
QY 1093 TCGCAAAAGAACACGAGAGGCACTCCCGCTGTCTGTATATATGGAAGTGTGGCGGTTCGCC 1152
Db TCGGCAAGGCGCGCTGGAG-CGCGCGCTGAAGAGCGGCAAGCAACGCGCATGTTCA-- 501
QY 1153 TGTGTTCTGACATATCTGTCCCGCAGGAATGTGAGCGCATGTGTGTGTAAGTTGGCGCTG 1212
Db ---TGTTCGACATCGCGCTGCCCGGACATCGAGCGGAAAGTTCGGCGAGCTGGACGAG 444

Qy 1213 CACGAGTGATACATGTGACGACTTGAAGAGGTGGTGGAGCCAAACAGAGAGACCGGC 1272
Db 443 TCTACCTCTATAGGTGACGACCTTCAACGAAGTAGTGGGGAGAAATCTCAAGAGCGGCC 384
Qy 1273 TCAGGAAGCAATGGAGCGCGAGACAATCATACCGAAGAACTGAGACGGTTTCGAGGCAT 1332
Db 383 AGGGCGCTCCAGCGCCCGAGGAACTGGTCCGACGGGGTCCGCGAGTTTCATGAGC 324
Qy 1333 GGAGGAGCTCGCTGGAGACCGTTCCGACCATCAAGAAAGCTGAGTGTCTGACCGGACAGGA 1392
Db 323 GCCTGCGGAACTGGCGCGGTGACGCTGTGCGCGCTATCGCGCAGCAGCGCGAGCGCC 264
Qy 1393 TCAGGGCTCGGAGCTCGAGAGTGCCTGAGAAAGTAGTGGAGCGCCCTCACCAAGA 1452
Db 263 TGCAGATGAGAGCTGGCAGGCCCGAGCGCACTGGCCAGCGGTGGCGACCCCGCGC 204
Qy 1453 AGATGAGGAGGCGCATCGAGAGCTGAGCACCGGCATCGTTAAACAGCTCTCTCCATGGCC 1512
Db 203 AGGTG-----ATGCGCCAGCTGGCCCGCGGCTGACCAACAACTGCTGCACGGC 153
Qy 1513 CGCTGACGACCTGAGTGGCGAGCGGAGCGACCGCACCTTTGACGAGACGCTCG 1569
Db 152 CCAGCGTGCAGATGAAGAAATGTCCCGCGAGGCGCGCATCGACGCGCTGGCCCTGG 96

RESULT 3

US-09-252-991A-14075
; Sequence 14075, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14075
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14075

Query Match 4.3%; Score 82.6; DB 4; Length 1026;
Best Local Similarity 49.6%; Pred. No. 2.2e-09;
Matches 326; Conservative 0; Mismatches 319; Indels 12; Gaps 4;
Qy 313 CCATAGCTGTAGTGGCTCAGTGTAACACAGCACCGAGTGGAGATGCGTGAAGAACTTG 372
Db 141 CTTTCATTTGCCCTCGGCATCAATCAAGACCGGCTCGGTGGTGTCCGCGAGCGGTGG 200
Qy 373 CTGTTCAGAGGAACCTGCGCCCGTCTATTCAAGAACTCACTAGCCTGAACCATATTG 432
Db 201 CTTTCACTCCGAAACAGATGGTGAAGCGCTGAGAGCTCTGCGCGCTGACCAAGCC 260
Qy 433 AAGAGGCTGTGTTCTTAGTACCTGTAATAGAAATGGAATTTATGTGGTGGCGCTCATCAT 492
Db 261 GCGAGGCGGCATCTCTGTGACCTGCAACCGAGCGAACTCTACCTGGAGAT-CGACCAT 319
Qy 493 GGAACGCTGTATCAGAGAAAGTAGTGAGTGGATGCGAAGAAAGTGGTATTCCCGCTT 552
Db 320 CCAGCGCGGAC-----GAGTGTGCTGCTGGCTGGCGGACTACCATCGCTGACTCTCG 374
Qy 553 CCGAGCTTAGGAGACCTGTTTCATCTTGCAGAGCAGTGTGACCAAGCGCATCTGTTTG 612
Db 375 ACAGATTGCGCGCTGTGGCTATGTGCACCGAGACGAGCGGGTGCAGCATGATGC 434
Qy 613 AGGTGTGAGTGGCCTTGACTCTTTGTTGTTCTCGTGAAGGACAAATCCTTTGCTCAGGTTA 672

Db 435 GCGTGGCTCCGGCTCGACTCGANTGTCTCGCGAGCCGCGAGATCTCGGCCAGATGA 494
Qy 673 AACAAAGTTGTGAGAGTGGACAGAAACAGTGGAGGCTTGGGAAAGAACATCGATAGGATGT 732
Db 495 AGTGGGCTACGCGCTCGCGCGGAGCGCGGACCGTCCGTCCGCTGCTCGCGCGCTGT 554
Qy 733 TCAAGATGCAATCAGTCTGGAAAGGTTGTCCGACGCGAGACCAACATATCATCTGGTG 792
Db 555 TCCAGGCGACCTTTCAGTACGCGGAGACCGTCCGACCCGATACCGCATCGCGGAGAAC 614
Qy 793 CTGTTTCTGTGAGTTCAGCGCGGTTGAACTGGCCCTGATGAAGCTTCCGAAAGTCTGAAG 852
Db 615 CGGTGTCGTGGGTTCCGCGCGGTGAGCTTGGCCAAAGCAGAT-CTTCAGCGACCTGCAC 673
Qy 853 CACTGTGAGTGTGCTCTGATTTGCTGTGATAAATGGGAAAGCTAGTGTATCAAAAC 912
Db 674 GCGAGCCA-----GGCGCTGCTGATCGCGCGCGGAGACCATCACCTGTGTGCGCGCC 728
Qy 913 ATCTGGTTCGCAAGGATGCAAGAGGTTGTTGTGTGAACCGCTCCGTGGAAAGGG 969
Db 729 ACCTGTTTCGAGCAAGCGGTGAAGCGCATCGTGTGGTGGCCAAACCGCACCTCGAGCGC 785

RESULT 4

US-09-134-001C-218
; Sequence 218, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 218
; LENGTH: 1347
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-218

Query Match 3.4%; Score 64.6; DB 4; Length 1347;
Best Local Similarity 45.4%; Pred. No. 2.5e-05;
Matches 232; Conservative 0; Mismatches 279; Indels 0; Gaps 0;
Qy 316 TAGCTGTGATAGGCCTCAGTGTACACAGCACCGAGTGGAGATCCGCTGAAGAACTTGTCTG 375
Db 8 TTGTTGCAATTAGCATAAATCATCGAACAGCTGATGTAAACATTAAGAGAGCAAGTTGCTT 67
Qy 376 TTGACAGGAGACTGTGGCCCCGCTATTCAAGAACTCACTAGCTGAAACCATATTGAAG 435
Db 68 TTAGAGATGATGCTCTACGATTAGCACATGAAGATTTATATGAACATAAAGCAATTTTAG 127
Qy 436 AGGCTGCTGTTCTTAGTACTCTGTAATAGAAATGAAATTTATGTGGTGGCGCTCATGGA 495
Db 128 ABAATGCTATTTTATCTACATGTAACTGTAAGTATACGCTATTGTTGATCAAGTTC 187
Qy 496 ACGTGGTATCAGAGAAAGTAGTGGACTGGATGTCGAAGAAAGTGGTATTCCCGCTTCCG 555
Db 188 ATACAGGACGTTATTATATACAAAGATTTTATAGCGGCTCTTTTGGATTTGAGGTAGATG 247
Qy 556 AGCTTAGGAGGACCTGTTTCATCTTCGGAAGCAGTGTGCCACACCGCATCTCTGTTGAGG 615
Db 248 ATATTAAAGATATGTCGGAAGTTAAAGTGGGGACGATGCAAGTTGAACATTTATTGCGTG 307
Qy 616 TGTACAGCTGCGCTTGACTCTTGGTTCTCGGTGAAGGACAAATCCTTGTCTCAGGTTAAAC 675
Db 308 TCACTTCTGGCTTAGATTCAATTGCTGTTGGTGAACACAAATTTTAGGACAAATCGCGG 367

Qy	676	AGTTCTGAGGAGTGGACAGACACGTGGAGGCTTGGGAAAGAACATCGATAGAGTGTCA	735
Db	368	ATGCATTTTCTTAGCGCAAAATCTGGTCAACTGGAAACGATTTTTTAATCATTTATTTA	427
Qy	736	AGGATCGATCACTGCTGGAAGAGCGTGTCCGAGGAGAGCAACATATCATCTGTGGTGTG	795
Db	428	AACAAGCGATTACTTTTGTCTAAAAAGCACACAGTGAACAGACATTTGCAGATAATGCTG	487
Qy	796	TTTCTCTCAGTTTCAGCGCGGTTGAACCTGGC	826
Db	488	TGAGTGTGTTCTTTATGCTGCTGTGTTGAATTAGC	518

RESULT 5

```

US-08-916-421B-1/c
: Sequence 1, Application US/08916421B
: Patent No. 6503729
: GENERAL INFORMATION:
: APPLICANT: Bult et al.
: TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
: Patent No. 6503729
: TITLE OF INVENTION: jannaschii
: FILE REFERENCE: PB275
: CURRENT APPLICATION NUMBER: US/08/916,421B
: CURRENT FILING DATE: 1997-08-22
: PRIOR APPLICATION NUMBER: US 60/024,428
: PRIOR FILING DATE: 1996-08-22
: NUMBER OF SEQ ID NOS: 3
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 1
: LENGTH: 1664976
: TYPE: DNA
: ORGANISM: Methanococcus jannaschii
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: (28222)..(28222)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc feature
: LOCATION: (28257)..(28258)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc feature
: LOCATION: (84773)..(84773)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc feature
: LOCATION: (84808)..(84808)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc feature
: LOCATION: (84812)..(84812)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc feature
: LOCATION: (98120)..(98120)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc feature
: LOCATION: (98159)..(98159)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc feature
: LOCATION: (98239)..(98239)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc feature
: LOCATION: (98266)..(98266)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc feature
: LOCATION: (98343)..(98343)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc feature
: LOCATION: (103998)..(103998)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc feature
: LOCATION: (148948)..(148948)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc feature
: LOCATION: (163385)..(163385)

```

NAME/KEY: misc feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1664855)..(1664855)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

Query Match 3.3%; Score 64.2; DB 4; Length 1664976;
Best Local Similarity 54.4%; Pred. No. 0.00036;
Matches 129; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
QY 590 TGATGCCACAGCCATCTGTTGAGGTGTCAGCTGGCCTTGACTCTTTGGTTCTCGGTGA 649
Db 1331928 TAAAGCAATAGAACATCTTTTATAGATTGTCATGTGGTTAGAGTCCATGATTGTTGGAGA 1331869
QY 650 AGGACAAATCTTGCTCAGGTGTAACAAAGTTGTGAGGAGTGGACAGAACAGTGGAGCCTT 709
Db 1331868 AGACCAATACTTGGGAGTTAAATAATGCTATCTAAAGCAAAAGAAAAGCAGAAAT 1331809
QY 710 GGGAAAGAACATCGATAGGATGTTCAAGATGCAATCACTGCTGGAAAGCGTGTCCGCGAG 769
Db 1331808 ATCCAAAAATCGAGAAAAATTTTAAAGGCAATACATCTGGACAAAGGCGCAAGAT 1331749
QY 770 CGAGCAACATATCATCTGCTGCTTTCTGTCAGTTTCAGCGCGGTTGAACTGGC 826
Db 1331748 AGAGACAAAGATAATGAGCGCGGGGTTTCAATTGGCTCTGCGCGAGTTGAAATTGGC 1331692

RESULT 6
US-09-328-352-517
; Sequence 517, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:

APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 517
LENGTH: 1290
TYPE: DNA
ORGANISM: Acinetobacter baumannii
US-09-328-352-517

Query Match 3.0%; Score 57.4; DB 4; Length 1290;
Best Local Similarity 45.2%; Pred. No. 0.001;
Matches 309; Conservative 0; Mismatches 356; Indels 18; Gaps 2;
QY 320 TGTGATAGGGCTCAGTGTACACACAGCACCAGTGGAGATCGTGCAAAAACCTTGTGTTGC 379
Db 18 TGCATTGGGTGTCAACCATCAACAGCTTCTGTAGAACTCCGCGAGCAATGCTTTCAA 77
QY 380 AGAGAACTGTGGCCCGTCTATTCAAGAACTCACTAGCCTGAACCATATTGAAGAGC 439
Db 78 TGCAGAGCGATTAAATAATTGCTTCCGAAACAGCTCATCAGAGTCTTTAAAAGACTT 137
QY 440 TGTGTTCTTAGTACCTGTAATAGAAATTTATGTGTGGCGCTATCATCGAACCG 499
Db 138 AGTGGTTGCTCAACGCTGTAACCGTACAGAACTGTACGCCATGCGGAA----- 186
QY 500 TGGTATCAGAGAAAGTAGTGACTGGATGTCGAAGAAAAGTGGTATCCCGCTCCGAGCT 559
Db 187 -GATGCCGAAAGCTTCTGAATGGTTAGCTGATGCCAATAATATGATGAAGAGTGT 245
QY 560 TAGGGAGCACCCTGTTTCATCTTGGCAAGCAGTGTGCCACAGCCTCTGTTTGAAGTGT 619
Db 246 AATTCATCATGTTTATCGTTACGAAATGCTCAAGCCATCACACATTTAATGCGTGTAGC 305
QY 620 AGCTGGCTTGACTCTTTGGTTCTCGGTGAAGCAAAATCTTCTCAGGTAAACAAAGT 679
Db 306 AAGTGGTTTGGACTCACTCATGTTGGGCGAAACCGCAAAATTTTAGGGCAGGTGAAAGTGC 365
QY 680 TGTGAGGAGTGGACAGAAACAGTGGAGGCTTGGGAAAGAAACATCGATAGGATGTTCAAGGA 739
Db 366 TCTAGCATTTCTTAAGAGAGCGCAACAGTTTCTCCAGAAATTAATATGATGATTTAGATA 425
QY 740 TGAATCACTGCTGGAAAGCGTGTCCGACGAGACCAACATATCATCTGTTGCTGTTTC 799
Db 426 CGCTTTTACGCGCTAAACGCTGTGCGCTCTGAAACAGCTGTAGGTAGTCATGCGTTTC 485
QY 800 TGTCACTTACGCGCGGTGAACTGGCCCTGTATGAAGCTTCCGAAAGTCTGAAGCAGTGT 859
Db 486 AATGGGTTATGCAAGTTGCCAGCTCGCTTTTACAGGTTTGTAGTAAACCTGAAAACTAAC 545
QY 860 AGCTAGGATGCTTCTGATTTGCTGTAATAAAGGAAAGCTAGTGATCAACATCTGGT 919
Db 546 GGTAA-----TGTGTTGCTGCTGCGAAATGAACAGCCTTGTGCCAAGCAATTTAGC 599
QY 920 TGCCAAAGGATCAAGAAAGTGTGTTGTGTGAACCGCTCCGTGGAAGGCGTGAATGCTAT 979
Db 600 TGAATAGGGCGTCCGCAAGATGATTATTTGTATCGTAGCCGTGAAGCTGCCGATCAATT 659
QY 980 TCGTAGGAGATCAAGATATATAG 1002
Db 660 GGCCCAAGAAATTCACATCAAG 682

RESULT 7
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.

```
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232.463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935.313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9pt-F18
; US-08-232-463-14
```

```
Query Match 2.9%; Score 56.4; DB 1; Length 7218;
Best Local Similarity 2.7%; Pred. No. 0.0031;
Matches 9; Conservative 199; Mismatches 120; Indels 0; Gaps 0;

Qy 1176 AGGAATGTCAGCGCATGTGTCTGAAGTTGGCGCTCCAGAGTGTACAATGTCAGCAG 1235
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1382 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1323
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1236 TTGAAGAGGTGGTGAAGCAACAGAGGACCGGCTCAGGAAGCAATGGAGGCGCAG 1295
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1222 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1263
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1296 ACATCATCACCGAAGCACTGAGACGGTTTCAGGCGATCGAGGAGCTCGCTGGAGACGTT 1355
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1262 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1203
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1356 CCGACCATCAAGAGCTGAGTCTGTCAGCGGACAGGATCAGGCGCTCGGAGCTCGAAG 1415
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1202 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1143
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1416 TGCTCGAGAAAGTAGGTGAGGACGCTCACCAGAGATGAGGAGACCATCGAGGAG 1475
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1142 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1083
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1476 CTGAGCAGCGGCATCGTTAAAGCTCC 1503
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1082 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1055
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 8
US-09-103-840A-2
```

```
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCES: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
; US-09-103-840A-2

Query Match 2.8%; Score 53; DB 3; Length 4403765;
Best Local Similarity 45.1%; Pred. No. 0.16;
Matches 243; Conservative 0; Mismatches 290; Indels 6; Gaps 1;

Qy 435 GAGCGTCTGTTCTTAGTACTGTAAATAGAAATGGAATTTATGTGGTGGCGCTATCATGG 494
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 601925 GAGCGATGTGTGTGCTGCTGCACTTGCACCGGCTGAGGCTCTACGCCGTAGTGAGCGGCTC 601984
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 495 AACCGTGGTATCAGAGAAGTAGTGGACTGTGGAAGAAAGTGTGATTTCCCGCTTCC 554
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 601985 CATGGCGGCTGTGCGGTGATCGGCGAGTCTTCCGCAACTCCGCTATGTCGATGGG 602044
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 555 GAGCTTAGGAGCACCTGTTTCATCTTGGGAAGCAGTATGCCACACCCCATCTCTTTGAG 614
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 602045 GAGCTGACCAAGTACGCATATGTCGCTACAGCGAGGACGAGTTGAGCACCTGTTCCGC 602104
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 615 GTGTGAGCTGGCGCTTGACTCTTTGGTTCTCGGTGAAGGACAAATCTTGTCTCAGGTTAAA 674
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 602105 GTTCCAGCGGCTTGACTCGGCGGTGATCGGCGAGCAGCAGGTGCTTGTGTCAGTGGCG 602164
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 675 CAAGTTGTGAGGAGTGACAGAACAGTGGAGGCTTGGGAAGAACATCGATAGGATGTC 734
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 602165 CGCGCTATCCGTCGCGGAATCCAACCGCACGTCGCGCGCTGTCACGAATTGGCC 602224
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 735 AAGATGCAATCACTGCTGGAAGCGTGTCCGAGCGAGACCAACATATCATCTGCTGCT 794
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 602225 CAGCGGCGCTGTGCGGTGGCAAGCAGTGCATCTCCGAAACCGCATTTGAGCTGCGGT 602284
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 795 GTTCTGTGAGTTGAGCGGCGGTTGAACCTGGCCCTGATGAAGCTTCCGAAAGTCTGAAGCA 854
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 602285 GCCTCGTGTGTGCTGCGCTGGGATGCGCGAGCGCAANTTGGGCTGCTGTCGCGGCG 602344
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 855 CTGTGAGTAGGATGCTTCTGATTTGGTGTGTTAAATGGGAAAGCTAGTGTATCAACAT 914
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 602345 ACAGCGC-----GGTGTGATCGCGCGCGCGATGCGCGCTGTGCGCGGTACAT 602398
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 915 CTGTTGCCAAAGATGCAAGAGGTTGTTGTTGTAACCGCTCCGCTGGAAAGGGTGA 973
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 602399 CTGACCGTCCCGGCTGCGGCGACATTCAGGTGCTCAACCGGTGTTGTCCTCCGCGCGCA 602457
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 9
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
```

APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
TUBERCULOSIS
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 4411529
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match
Best Local Similarity 2.8%; Score 53; DB 3; Length 4411529;
Matches 243; Conservative 0; Mismatches 290; Indels 6; Gaps 1;

QY 435 GAGGCTGCTGTTCTTAGTACCTGTAATAGATGGAATTTATGTCGGTGGCTATCATGG 494
DB 600565 GAGGGGATGGTCTGTGACCTTGCAACCGCTGAGGTCTACGCCGTAGTGACCGGTTC 600624
QY 495 AACCGTGTATCAGAGAAGTAGTGACCTGGATGTCGAAGAAAGTGGTATTCCTCCCTTCC 554
DB 600625 CATGGCGCCTGTCGGTGATCGGCGAGGTGCTTGCCGAACACTCCGGTATGTCGATGGG 600684
QY 555 GAGCTTAGGGAGCACCTGTTCTATCTTGGAGCAGTATGCAACAGCCATCTGTTTGA 614
DB 600685 GAGCTGACCAAGTACGCATATGTCCGCTACAGCAGGCGAGCTTGAGCACCCTGTCGCG 600744
QY 615 GTGTCAGCTGGCTTGACTTTTGGTCTCGGTGAAGCAGCAATCTTCTGCTCAGGTTAA 674
DB 600745 GTTGCACGCGCTGGACTCGCGGTGATCGCGAGCAGAGTCTTGGTCAGGTGCGC 600804
QY 675 CAAGTTGTGAGGAGTGGACAGAACAGTGGAGGCTTGGGAAAGAACATCGATAGGATGTC 734
DB 600805 CGCGCTATGCGTGCAGCAATCCAAACCGCAGCTGCGCGGTGTCACGAATTGGCC 600864
QY 735 AAGATGCAATCACTGCTGGAAGCGTGTCCGACGAGCAGCAATCATCTGTTGCT 794
DB 600865 CAGCGGCGCTGTGCGTGGGCAAGCGAGTGCACCTCCGAACCGCCATTTGACGCTGCCGT 600924
QY 795 GTTCTGTGAGTTCAGCGCGCTTGAATCGCCCTGATGAAGCTTCCGAAGTCTCAAGCA 854
DB 600925 GCCTCGTGGTGTGCTGCTGCTGCTGGAATGCGCGAGCCCAATTGGGCTGCTGGCGGC 600984
QY 855 CTGTCAGTAGGATGCTTCTGATTGCTGCTGTAATGGGAAAGCTAGTATCAACAT 914
DB 600985 ACGACCGC-----GGTGTGATCGCGCGCGCGGCGATGGCGGCTGTCGGCGGTACAT 601038
QY 915 CTGTTGCCAAGGATGCAAGAGTTGTTGTTGTAACCGCTCGTGGAAGGTTGGA 973
DB 601039 CTGACCCGCTGCGCGCTGCGGCACATTCAGGTGCTCAACCGTGTGTTGTCGCGCGCA 601097

RESULT 10
US-09-514-247A-9/c
Sequence 9, Application US/09514247A
Patent No. 6365361
GENERAL INFORMATION:
APPLICANT: TANABE SEIVAKU CO. LTD.
APPLICANT: TANIGUCHI, Tomoyasu
APPLICANT: MIZUKAMI, Junko
TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND ANTAGONIST TO PPA
FILE REFERENCE: TANIGUCHI=6
CURRENT APPLICATION NUMBER: US/09/514,247A
CURRENT FILING DATE: 2000-02-28
PRIOR APPLICATION NUMBER: PCT/JP98/03734
PRIOR FILING DATE: 1998-08-24
PRIOR APPLICATION NUMBER: JP231084/1997
PRIOR FILING DATE: 1997-08-27
NUMBER OF SEQ ID NOS: 10

SOFTWARE: PatentIn version 3.0
SEQ ID NO 9
LENGTH: 8147
TYPE: DNA
ORGANISM: human
FEATURE:
NAME/KEY: CDS
LOCATION: (819)..(8147)
US-09-514-247A-9

Query Match
Best Local Similarity 2.6%; Score 50; DB 4; Length 8147;
Matches 98; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 82 CGACGTACGACGACACCCCGCGGAGCAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 141
DB 295 CGCGAGCGGGGTGC 236
QY 142 CGTGGCCCTCTGCCAGAGGCTGGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 201
DB 235 CGCGCTGCGG 176
QY 202 CGCACGCGCGCGCGGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 259
DB 175 CG 118

RESULT 11
US-08-311-731A-128/c
Sequence 128, Application US/08311731A
Patent No. 6583266
GENERAL INFORMATION:
APPLICANT: SMITH, DOUGLAS
APPLICANT: MAO, JEN-I
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
RELATED TO MYCOBACTERIUM TUBERCULOSIS
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 411
CORRESPONDENCE ADDRESS:
ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
STREET: 600 ATLANTIC AVENUE
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,731A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: GATES, EDWARD R.
REGISTRATION NUMBER: 31,616
REFERENCE/DOCKET NUMBER: C0044/7125
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/720-3500
TELEFAX: 617/720-2441
INFORMATION FOR SEQ ID NO: 128:
SEQUENCE CHARACTERISTICS:
LENGTH: 42988 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: MYCOBACTERIUM LEPRAE
US-08-311-731A-128

Query Match 2.6%; Score 49.6; DB 4; Length 42988;
Best Local Similarity 45.4%; Pred. No. 0.19;
Matches 178; Conservative 0; Mismatches 214; Indels 0; Gaps 0;

QY 435 GAGGCTGCTGTTCTTAGTACCTGTAATAGAAATGGAATTTATGTGTCGCGCTATCATGG 494
DB 20252 GAGGCGATGGTCTGTGACCTGTAACTGTCGAGGTTTACCGCTGGTGAACGGTTC 20193

QY 495 AACCGTGGTATCAGAGAAGTAGTGGACTGGATGTCGAAGAAAAGTGGTATTCGCCGTTCC 554
DB 20192 CACGCGGATTATCAGTGATTGGCAAGTGTGCGGAATACTCCGCGATGTCGATCGG 20133

QY 555 GAGCTTAGGAGCACCTGTTCTATCTTGGCAAGCAGTGATGCCACAGCCATCTGTTTGA 614
DB 20132 GACCTGACCAATAACGCTTATGACCTACACGAGGCTGTGTCGAGCAGTGTGTCACG 20073

QY 615 GTGTACGCTGGCGCTTGAACCTTTTGGTTCTCGGTGAAGCAAAATCTTCTCAGGTTAA 674
DB 20072 GTTGCACCGGTCTGGATTACCGGTGTGTCGCGAGCAGCAGGTGCTGGTTCAGTGGT 20013

QY 675 CAAGTTGTGAGAGTGGACAGAACAGTGGAGGCTTGGGAAAGAAATCATGATAGGATGTC 734
DB 20012 CGCGGCTATGCGCGCGCTGAGGCTAATCGCACGCTAGGCCAGGTGTGTCATGAGGTGGCC 19953

QY 735 AAGGATGCAATCACTGCTGGAAGCGTGTCCGACGAGACCAATATCATCTGTTGCT 794
DB 19952 CAACGAGCGTTGTCGTTGGTGAAGCGGTGCAATTCGAAACTGCCATTAAGACGCTGCCGG 19893

QY 795 GTTCTGTGTCAGTTCAGCGCGCTTGAACCTGGC 826
DB 19892 GTCTCGGTGGTGTGCGTCCGCTTGGCATAGC 19861

RESULT 12
US-09-086-010-1/c
; Sequence 1, Application US/09086010
; Patent No. 6274338
; GENERAL INFORMATION:
; APPLICANT: Glimcher, Laurie H. et al.
; TITLE OF INVENTION: Human c-Maf Compositions and
; TITLE OF INVENTION: Methods of Use Thereof
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/086,010
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/030,579
; FILING DATE: 2-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Kara, Catherine J.
; REGISTRATION NUMBER: 41,106
; REFERENCE/DOCKET NUMBER: HUI-027CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1203 base pairs
; TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1203
US-09-086-010-1

Query Match 2.6%; Score 49.4; DB 3; Length 1203;
Best Local Similarity 51.6%; Pred. No. 0.061;
Matches 113; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 22 GAGGAGCTTGGGAAGTTGCCAAGGCTCCGATTTTCGCTAATGCGACGATAATGGCGACCA 81
DB 776 GAGAAGCGGTGCTGCAAGTGCAGGCGCGCGCGGTGTGGTGTGCGGCGCGCGCGCGCG 717

QY 82 CGAGCTCAGCGACCAACCCCGCGCAGCAGCGCCACCGCCACCGCCAGCGCGCGCGCGT 141
DB 716 GCGCCCGCGCGCTCCCG 657

QY 142 CGTCGCGCTCTGCGCAGAGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 201
DB 656 CG 597

QY 202 GCGAGCG 240
DB 596 GCG 558

RESULT 13
US-08-466-103A-3
; Sequence 3, Application US/08466103A
; Patent No. 5856124
; GENERAL INFORMATION:
; APPLICANT: Reppert, Steven M.
; APPLICANT: Ebisawa, Takashi
; TITLE OF INVENTION: HIGH-AFFINITY MELATONIN
; TITLE OF INVENTION: RECEPTORS AND USES THEREOF
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,103A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/319,887
; FILING DATE: 07-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/261,857
; FILING DATE: 17-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Frazer, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00786/250002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1149 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double


```

;
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 49...1146
; OTHER INFORMATION:
; US-08-466-103A-3

```

Query Match 2.5%; Score 48.8; DB 2; Length 1149;
Best Local Similarity 51.9%; Pred. No. 0.081;
Matches 110; Conservative 0; Mismatches 102; Indels 0

Matches	110;	Conservative	0;	Mismatches	102;	Indels	0;	Gaps	0;
Qy	79	CAACGAGCTCAGCGACCAACCGCGCCGCGCAGCGCGCCACCAACGCGCAACGCGCGGGGT	138						
Db	8	CGACGCTCTGGGGATCCACCGGGCCCGCCCTGCACGCGATGCGGGGCGGCTGTGGG	67						
Qy	139	CGTGTGTCGGCCCTCTGCCAGAGAGGTGGCCGGCGGGCGGACGCGCGCTCCGGGGTGTGTC	198						
Db	68	GCTCGCGCGGGCGGGACCCCCAAGGGCAACGGCAGCAGCGGCTGCTCAACGTCTCGCAGG	127						
Qy	199	GGTGCGACGCGCGCGCGCTGCGAGGCCACAGCGCGAGGCGCTGGCAAGGCTGCCAGCGTCG	258						
Db	128	CGCGCGCGCGCGCGGGACGGTGTGGGCGCGGGCCCTCGTGGCTGGCCGCCACCTCG	187						
Qy	259	CGCGCCCTCGAGCAGTTCAAGATATCCGCCGAC	290						
Db	108	CCTCCATCCTCATCTTCCACCATCGTGGTGAC	219						

RESULT 14

US-09-410-551B-1/C
; Patent 1, Application US/09410551B
; Sequence No. 6503737
; GENERAL INFORMATION:
; APPLICANT: KOSAN BIOSCIENCES, Inc.
; APPLICANT: REEVES, CHRISTOPHER
; APPLICANT: CHO, DANIEL
; APPLICANT: KHOSLA, CHAITAN
; APPLICANT: SANTI, DANIEL
; APPLICANT: WU, KAI

TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES AND RECOMBINANT DNA

```

: TITLE OF INVENTION: CONSTRUCTS THEREFOR
: FILE REFERENCE: 3062-20026.00
: CURRENT APPLICATION NUMBER: US/09/410,551B
: CURRENT FILING DATE: 1999-10-01
: PRIOR APPLICATION NUMBER: US 60/139,650
: PRIOR FILING DATE: 1999-06-17
: PRIOR APPLICATION NUMBER: US 60/123,810
: PRIOR FILING DATE: 1999-03-11
: PRIOR APPLICATION NUMBER: US 60/102,748
: PRIOR FILING DATE: 1998-10-02
: NUMBER OF SEQ. ID NOS: 72

```

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 77536

; TYPE: DNA

ORGANISM: *Streptomyces hygroscopicus*

FEATURE: WAMP/WEB

```

; NAME/KEY: CDS
: LOCATION: (533'15) (71465)

```

US-09-410-551B-1

Query Match

Query Match 2.5%; Score 48.4; DB 4; Length 77536;
Best Local Similarity 53.1%; Pred. No. 0.42;
Matches 103; Conservative 0; Mismatches 91; Indels 0;

	Matches	103;	Conservative	0;	Mismatches	91;	Indels	0;	Gaps	0;
Qy	97	CGCGCCGCGCAGCGCGCACACCGCCAGCGCGGGGCGTCTGTCGCCCTCTGCC	156							
Db	29025	CGCGCGGCGCACACCTTGCACGCGATCGCGGGCTCTCGCGGGAGCGCGAAA	28966							

Qy 157 AGAGGTGCGCCGGCGGCACGGCGCGTTCGGGGTGTCGGTGCGACGCCCGCGCGC 216
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Dδ 28965 AGGACGTGCTCGCCCTCGTCCGGCAGTGCAACCGCGCCGCTACTCGGCACACACGGTGCCG 28906

217 TGAGGCCCAGCGCAGCGCGTGGCAAGGCTGCAGCGTCGCGCCCTCGAGCAGTTCA 276
28905 CACGGGTGAGCGCAGCCACGTTCAAGGAACTCGCGCTCGACTCGCTCATGGCGATCC 28846
277 AGATATCGCGCCGAC 290
28845 GGCTGCGCAACGCC 28832

RESULT 15

```

US-09-103-840A-2/c
; Patent 2, Application US/09103840A
; Sequence No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLETSCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR
; IDENTIFICATION OF TUBERCULOSIS
; FILE REFERENCE: 24366-20007, 00
; CURRENT APPLICATION NUMBER: US/09/103-840A-2
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at variable positions
; -09-103-840A-2
; US-09-103-840A-2

```

Query Match

Query Match	2.5%;	Score	47.6;	DB	3;	Length	4403765;
Best Local Similarity	50.9%;	Pred.	No.	2.5;			
Matches	113;	Conservative	0;	Mismatch	109;	Indels	0; Gaps 0;
QY	46	CCTCCGATTTCGCTAATCGACGATAATGGCGACACACGACGTCAGCGACCCACCGCCGCCG	105				
Db	3793567	CCACCGCGTTGTTGAGCCCGCGCGCGCCCGGACCGCGCGCGCGCTGCCGTAG	3793508				
QY	106	GAGCAGCCGCCACCAACCGGCAAGCCGCGGGGTGCTGCTGGCCCTTCCACAGAGGTGG	165				
Db	3793507	AGCCACCCGCGCGCCCGCGCGCCCGCGGTCCCTCCCGTGCCTCCGCTAGGGTGGCG	3793448				
QY	166	CCGCGCGGCGCAGCGCGGCGCTCCGCGGGTGTGCGGTGCGACCCCGCGCGGTGGAGCCC	225				
Db	3793447	CCACGCTGGAATTTCTTCCGCGCGGCCCGCGCGTCCCGCGCGCCGTCATGCCGACC	3793388				
QY	226	AGGCGCAGGCCGTGGCAAGGCTGCACGCTGCCGCCCTCG	267				
Db	3793387	AGGTACCGCGCGCCGACACCGCGCGCGCGCGCGCGCGCTTG	3793146				

Search completed: December 15, 2003, 12:48:39
Job time : 139 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 12:53:43 ; Search time 39 seconds
(without alignments)
3513.485 Million cell updates/sec

Title: US-10-018-902-4
Perfect score: 2606
Sequence: 1 MATTTSATTAATAAATTAKP.....DWEKALIEQIKAKVEKTON 531

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 23:.*
1: sp archea:.*
2: sp bacteria:.*
3: sp fungi:.*
4: sp human:.*
5: sp invertebrate:.*
6: sp mammal:.*
7: sp mhc:.*
8: sp organelle:.*
9: sp phage:.*
10: sp plant:.*
11: sp rodent:.*
12: sp virus:.*
13: sp vertebrate:.*
14: sp unclassified:.*
15: sp virus:.*
16: sp bacteriap:.*
17: sp archaeap:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2393.5	91.8	536	Q9FW00	Q9fw00 oryza sativ
2	1789	68.6	542	Q9ZPK4	Q9zpk4 glycine max
3	1440	55.3	339	Q9FVS1	Q9fyv1 hordeum vul
4	1334	51.2	524	Q9SUX1	Q9sux1 arabidopsis
5	1326.5	50.9	522	Q9FPR7	Q9fpr7 chlamydomon
6	1063	40.8	252	Q9FVS2	Q9fyv2 hordeum vul
7	910.5	34.9	248	Q9AVP6	Q9avt6 amaranthus
8	893.5	34.3	426	Q8DI53	Q8dif5 synechococc
9	646	24.8	419	Q8DFF8	Q8dff8 vibrio vuln
10	618.5	23.7	446	Q8GCB0	Q8gcb0 bacillus me
11	597	22.9	416	Q8EAR2	Q8ear2 shewanella
12	588	22.6	457	Q8EPM7	Q8epm7 oceanobacil
13	570.5	21.9	484	Q8FT03	Q8ftf3 escherichia
14	543.5	20.9	448	Q8CNY6	Q8cny6 staphylococ
15	488.5	18.7	463	Q8GMZ2	Q8gmz2 corynebacte
16	483.5	18.6	463	Q8NT91	Q8nt91 corynebacte

17	465.5	17.9	466	16	Q8FSP8	Q8fsf8 corynebacte
18	451	17.3	460	17	Q8TTG0	Q8ttg0 methanosarc
19	421	16.2	420	16	Q8D2K8	Q8d2k8 wiggleswort
20	294	11.3	329	16	Q8R687	Q8r687 fusobacteri
21	287	11.0	56	10	Q8RUB4	Q8rub4 zea mays (m
22	282	10.8	55	10	Q8S2X3	Q8s2x3 zea mays (m
23	277	10.6	56	10	Q8LNE9	Q8lne9 oryza sativ
24	256.5	9.8	377	2	Q9JA66	Q9ja66 uncultured
25	190	7.3	184	2	Q9AQ47	Q9aq47 selenomonas
26	156.5	6.0	644	16	Q8RAZ5	Q8raz5 thermomanae
27	152	5.8	807	10	Q48724	Q48724 arabidopsis
28	152	5.8	816	5	Q8IU20	Q8iu20 branchiosto
29	150	5.8	794	17	Q8ZVH4	Q8zv4 pyrobaculum
30	148	5.7	853	10	Q9LFE4	Q9lfe4 arabidopsis
31	147.5	5.7	739	17	Q8TUP6	Q8tup6 methanosarc
32	147.5	5.7	791	17	Q9YB89	Q9y89 aeropyrum p
33	147.5	5.7	1307	16	Q8K792	Q8k792 streptococc
34	147	5.6	530	16	Q9XOM7	Q9xom7 thermotoga
35	144.5	5.5	1751	5	Q8T0G9	Q8t0g9 plasmodium
36	143.5	5.5	599	2	Q44107	Q44107 anaplasm m
37	143.5	5.5	779	10	Q9SZB6	Q9szb6 arabidopsis
38	143.5	5.5	817	10	Q8LHQ6	Q8lhq6 oryza sativ
39	143.5	5.5	1307	16	Q8P1P0	Q8p1p0 streptococc
40	142	5.4	2158	10	Q9LUT5	Q9lut5 arabidopsis
41	140.5	5.4	1313	10	Q9XIP6	Q9xip6 arabidopsis
42	140	5.4	1963	5	O02244	O02244 caenorhabdi
43	139.5	5.4	1728	10	Q9LUI2	Q9lui2 arabidopsis
44	139	5.3	1300	4	Q94986	Q94986 homo sapien
45	139	5.3	1978	4	O15154	O15154 homo sapien

ALIGNMENTS

RESULT 1

Q9FW00 PRELIMINARY; PRT; 536 AA.
ID Q9FW00
AC Q9FW00;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Putative glutamyl-tRNA reductase (EC 1.2.1.-) (GLUTR).
GN OSUNBB0073N24.1.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Erechtoidae; Oryzae; Oryza.
OX NCBI_taxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Jenkins C.N., Burr P.C.,
RA Hsiao J., Zismann V., Pai G., Bowman C.L., Fujii C.Y., VanAken S.E.,
RA Bowman C.L., Craven B., Uterback T.R., Khalak H., Feldblyum T.V.,
RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 10 BAC OSUNBB0073N24 genomic sequence.";
RL submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: GLUTAMYL-TRNA (GLU) + NADPH = GLUTAMATE-1-
SEMIALDEHYDE + NADP(+) + TRNA (GLU).
CC -1- PATHWAY: PORPHYRIN BIOSYNTHESIS BY THE C5 PATHWAY, FIRST STEP.
CC -1- SIMILARITY: BELONGS TO THE GLUTAMYL-TRNA REDUCTASE FAMILY.
DR EMBL; AC078840; AAG13620.1; -.
DR Gramene; Q9FW00; -.
DR InterPro; IPR000343; Glutr.
DR Pfam; PF00745; Glutr; 1.
DR TIGRfam; TIGR01035; hema; 1.
DR PROSITE; PS00747; Glutr; 1.
DR NADP; Oxidoreductase; Porphyrin biosynthesis.
SQ SEQUENCE 536 AA; 58288 MW; 4C1C7E4A58E3C640 CRC64;

Query Match 91.8%; Score 2393.5; DB 10; Length 536;
Best Local Similarity 91.6%; Pred.No. 3.le-132;
Matches 492, Conservative 15; Mismatches 23; Indels 7; Gaps 4;

```

QY 1 MATTTTATATAAATAATKPRGSSA---LQORVAGGRRRSG-VVRCDAAQV--AAQ 54
DB 1 MASTTATAAGGAFAAATKTRAGSSAAGGACARVAAGRRRSGVVVRCD-AGVEAQAQ 59
QY 55 AVAKAASVAALAEQFKISADRYMKERSTTAVIGLSVHTAPVEMREKLAVALPRAIOEL 114
DB 60 AVAKAASVAALAEQFKISADRYMKERSTTAVIGLSVHTAPVEMREKLAVALPRAISEL 119
QY 115 TSLNHIIEAAVLSTCNRMIEIYVVALSWNRGIREVVDMWMSKSGIPASELREHLFIRSSD 174
DB 120 TSLNHIIEAAVLSTCNRMIEIYVVALSWNRGIREVVDMWMSKSGIPASELREHLFIRSSD 179
QY 175 ATRHLFEVSAGLDSVLGEGQILAQVKQVRSQNSGGLGKNIIDRMFKDAITAGKRVSE 234
DB 180 ATRHLFEVSAGLDSVLGEGQILAQVKQVRSQNSGGLGKNIIDRMFKDAITAGKRVSE 239
QY 235 TNISSGAVSVSSAAVELALMKLPKSEALSARMLLIGAGKMGKLVKHLVAKGCKVWVYN 294
DB 240 TNISSGAVSVSSAAVELALMKLPKSEALSARMLLIGAGKMGKLVKHLVAKGCKVWVYN 299
QY 295 RSVERDAIREEMKDIETVYRPLSDMYQAAAEADVFTSTASSETSLFAKEHAELPPVSD 354
DB 300 RSVERDAIREEMKDIETVYRPLSDMYQAAAEADVFTSTASSETSLFAKEHAELPPVSD 359
QY 355 TMGGVRLFVDISVPRNVACSEVGAARVYVDDLKEVVEANKEDRLKAMEAQTIITEE 414
DB 360 AMGGVRLFVDISVPRNVACSEVGAARVYVDDLKEVVEANKEDRLKAMEAQTIITEE 419
QY 415 LRRFEAWRDSLETVTIKKLSYADRIASELEKCLQKVGEDALTKKMRRAIEELSTGIV 474
DB 420 LRRFEAWRDSLETVTIKKLSYADRIASELEKCLQKVGEDALTKKMRRAIEELSTGIV 479
QY 475 NKLLHGPLOHRCDSGDSRTLDTELNNHMLNRMFSLDMEKAIIEQKIKAKVEKTON 531
DB 480 NKLLHGPLOHRCDSGDSRTLDTELNNHMLNRMFSLDMEKAIIEQKIKAKVEKTON 536

RESULT 2
Q9ZPK4 PRELIMINARY; PRT; 542 AA.
ID Q9ZPK4
AC Q9ZPK4;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAY-2003 (TREMBlrel. 23, Last annotation update)
DE Glutamyl-tRNA reductase (EC 1.2.1.-) (GluTr).
GN GTR1.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Essex;
RX MEDLINE=99137912; PubMed=9952455;
RA Sangwan I., O'Brian M.R.;
RT "Expression of a soybean gene encoding the tetrapyrrole-synthesis
RT enzyme glutamyl-tRNA reductase in symbiotic root nodules.";
RL Plant Physiol. 119:593-598 (1999).
CC -1- CATALYTIC ACTIVITY: GLUTAMYL-TRNA (GLU) + NADPH = GLUTAMATE-1-
CC SEMIALDEHYDE + NADP(+) + TRNA (GLU).
CC -1- SIMILARITY: BELONGS TO THE GLUTAMYL-TRNA REDUCTASE FAMILY.
CC EMBL; AF105221; AAD16897.1; -.
DR InterPro; IPR000343; Glutr.
DR InterPro; IPR000594; Thif domain.
DR Pfam; PF00745; Glutr; 1.
DR TIGRFAMs; TIGR01035; hema; 1.
DR PROSITE; PS00747; Glutr; 1.
DR KQ NADP; Oxidoreductase; porphyrin biosynthesis.
SQ SEQUENCE 542 AA; 59204 MW; CB85AA5494F09792 CRC64;

```

```

Query Match 68.6%; Score 1789; DB 10; Length 542;
Best Local Similarity 67.3%; Pred. No. 8e-97;
Matches 366; Conservative 70; Mismatches 88; Indels 20; Gaps 5;

QY 1 MATTTT-----ATTAAATAATKPRGSSSALCORVAGGRRRSGVVVRCD 46
DB 1 MAVTSPPGAKLEALLKCGSSNAATATATTTT---HLSCFKTRKTLVQSGRGPIREA 57
QY 47 AGV-BAQAQAVAKAASVAALAEQFKIS-ADRYMKERSTTAVIGLSVHTAPVEMREKLA 104
DB 58 SSASDVVADATKKAASVSALAEQFKIS-ADRYMKERSTTAVIGLSVHTAPVEMREKLA 117
QY 105 ELMPRATQELTSLNHIIEAAVLSTCNRMIEIYVVALSWNRGIREVVDMWMSKSGIPASE 164
DB 118 AEMPRATQELTSLNHIIEAAVLSTCNRMIEIYVVALSWNRGIREVVDMWMSKSGIPASE 177
QY 165 EHLFILRSSDATRHLFEVSAGLDSVLGEGQILAQVKQVRSQNSGGLGKNIIDRMFKDA 224
DB 178 QHQFLLYNKDATQHLFEVSAGLDSVLGEGQILAQVKQVRSQNSGGLGKNIIDRMFKDA 237
QY 225 ITAGKRVSETNISGAVSVSSAAVELALMKLPKSEALSARMLLIGAGKMGKLVKHLVA 284
DB 238 ITVGRVKTETNIAAGAVSVSSAAVELALMKLPKSEALSARMLLIGAGKMGKLVKHLVA 297
QY 285 KGCKVWVWNRSEVERDAIREEMKDIETVYRPLSDMYQAAAEADVFTSTASSETSLPAKE 344
DB 298 KGCTKMWVWNRSEVERDAIREEMKDIETVYRPLSDMYQAAAEADVFTSTASSETSLPAKE 357
QY 345 HAEALPPVSDTMGGVRLFVDISVPRNVACSEVGAARVYVDDLKEVVEANKEDRLKAME 404
DB 358 DVKELPATDEVGRRFLFVDISVPRNVACSEVGAARVYVDDLKEVVEANKEDRLKAME 417
QY 405 MEAQTIITEELRRFEAWRDSLETVTIKKLSYADRIASELEKCLQKVGEDALTKKMR 464
DB 418 MEAQTIITEELRRFEAWRDSLETVTIKKLSYADRIASELEKCLQKVGEDALTKKMR 476
QY 465 AIELSTGIVNKLHGPLOHRCDSGDSRTLDTELNNHMLNRMFSLDMEKAIIEQKIKAK 524
DB 477 AVDDLRSQVGNKLHGPLOHRCDSGDSRTLDTELNNHMLNRMFSLDMEKAIIEQKIKAK 536
QY 525 KVEK 528
DB 537 KVEQ 540

RESULT 3
Q9FYS1 PRELIMINARY; PRT; 339 AA.
ID Q9FYS1
AC Q9FYS1;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Glutamyl-tRNA reductase (EC 1.2.1.-) (GluTr) (Fragment).
GN HEMA2.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Bonus;
RA Hanson M.;
RT "Basic characterization of two barley hema promoter regions.";
RT Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: GLUTAMYL-TRNA (GLU) + NADPH = GLUTAMATE-1-
CC SEMIALDEHYDE + NADP(+) + TRNA (GLU).
CC -1- PATHWAY: PORPHYRIN BIOSYNTHESIS BY THE C5 PATHWAY; FIRST STEP.
CC -1- SIMILARITY: BELONGS TO THE GLUTAMYL-TRNA REDUCTASE FAMILY.
DR EMBL; AF294753; AAG02480.1; -.
DR InterPro; IPR000343; Glutr.
DR Pfam; PF00745; Glutr; 1.
DR PROSITE; PS00747; Glutr; 1.

```



```
Query Match      50.9%; Score 1326.5; DB 10; Length 522;
Best Local Similarity 53.7%; Pred. No. 9.4e-70;
Matches 275; Conservative 79; Mismatches 121; Indels 37; Gaps 5;

Qy 2 ATTTTATTAATAAATTAKPRGSSSALCORVAGGRRSGVVRCDAAAGVAAQAQAA 61
Db 34 AATQTATTAASSTTKLP-----ASHLE-----SS 59

Qy 62 VAALAEQFKISA-DRYM-KERSTIAVIGLSVHTAPVEMREKLAVALPRAIOELTSLNH 119
Db 60 KKALDSLKQAVNRVYAGDKSSITAGLTINAPVELREKLAVALPRAIEELCOFPH 119

Qy 120 IEAAVLSTCNRMELIYVVALSWNRGIREVVDWMSKSGIPASELREHLFILRSSDATRHL 179
Db 120 IEAAVLSTCNRMELIYVGLSWHRGVRVEEWSLRTSGVPLDELRLPVLFLLRDRDATHL 179

Qy 180 FEVSAGLDSLVLGEGQILAQVQVVRSGONGSGGLGKNIDRMFKDAITAGKVRSETNLS 239
Db 180 MRVSGGLDSLVMGEGQILAQVQVYKVGQNCFGFGRHLNGFQKAITAGKVRSETNLS 239

Qy 240 GAVSVSSAAVELALMKLPKSPALSGARMILLIAGKMGKLVIRHLVAKGCKKYVVYVNRSV 299
Db 240 GSVSVSSAAVELAQKLPKTHNWSDAKVCIIAGKMGSTLLVHLQSKGCKEYTVLNRSLPR 299

Qy 300 VDAIREENKDJIEIYRPLSDMYQAAAADVVFTSTASSETSLFAKEHAELPPVSDTMGGV 359
Db 300 AQALAEPEPKVFNHILMPDLQCCVEASDVIFAASGSEELIHKHEVEMSKPSDVVVGSK 359

Qy 360 RLFDVDSVPRNVSCVSVGAARVYVVDLKEVVEANKEDRLRKAMEAQTIITIELRFE 419
Db 360 RRFVDSVPRNIAPAINELEHGIYVVDLKEVVAANKEGRAQAAAEVLIRBEOQAFE 419

Qy 420 AWRDSLETVPYTKLRSYADRIRASELEKICQKQGEDALTKMRRRAIEELSTGVNKLH 479
Db 420 AWRDSLETVPYTKALRSKAETIRAAEPEKAVSRIGE-GLSKKQLKAVELSKGVNKLH 478

Qy 480 GPLQHLRCDGSDSRTLDETLENMHALNMFSL 511
Db 479 GPMTALRCDDGTPDAVGQTLNMEALERMFO 510

RESULT 6
Q9FY52 PRELIMINARY; PRT; 252 AA.
AC Q9FY52;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glutamyl-tRNA reductase (EC 1.2.1.-) (GluTR) (Fragment).
GN HEMAL.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Bonus;
RA Hansson M.;
RT "Basic characterization of two barley hema promoter regions.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: GLUTAMYL-TRNA(GLU) + NADPH = GLUTAMATE-1-
CC SEMIALDEHYDE + NADP(+) + TRNA(GLU).
CC -1- PATHWAY: PORPHYRIN BIOSYNTHESIS BY THE C5 PATHWAY; FIRST STEP.
CC DR EMBL; AF294752; AAG02479.1; -.
CC InterPro; IPR000343; GluTR.
DR Pfam; PF00745; GluTR; 1.
DR PROSITE; PS00747; GluTR; 1.
KW NADP; Oxidoreductase; Porphyrin biosynthesis.
FT NON TER 252
FT SEQUENCE 252 AA; 26654 MW; 4D3C0908EAAAF0FE CRC64;
```

```
Query Match      40.8%; Score 1063; DB 10; Length 252;
Best Local Similarity 86.7%; Pred. No. 9.1e-55;
Matches 222; Conservative 10; Mismatches 20; Indels 4; Gaps 2;

Qy 1 MATTTTATTAATAAATTAKPRGSSSALCORVAGGRRSGVVRCDAAAGVAAQAQAA 60
Db 1 MAGATSA-TAAGAFAAAKARGPAAACPMWVAAGRRRSGVVRCDAGG---DAQAASKAA 56

Qy 61 SVAALAEQFKISADRYMKERSTIAVIGLSVHTAPVEMREKLAVALPRAIOELTSLNH 120
Db 57 SITALEQFKISADRYMKESKSIIVIGLSVHTAPVEMREKLAVALPRAISELTSLNHI 116

Qy 121 EEAVALSTCNRMELIYVVALSWNRGIREVVDWMSKSGIPASELREHLFILRSSDATRHL 180
Db 117 EEAVALSTCNRMELIYVVALSWNRGIREVVDWMSKSGIPASELREHLFILRDSDATRHL 176

Qy 181 EVSAGLDSLVLGEGQILAQVQVVRSGONGSGGLGKNIDRMFKDAITAGKVRSETNLS 240
Db 177 EVSAGLDSLVLGEGQILAQVQVVRSGONGSGGLGKNIDRMFKDAITAGKVRSETNLS 236

Qy 241 AVSVSSAAVELALMKL 256
Db 237 AVSVSSAAVELAMMKL 252

RESULT 7
Q9AVF6 PRELIMINARY; PRT; 248 AA.
AC Q9AVF6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glutamyl-tRNA reductase (Fragment).
GN GLUTR.
OS Amaranthus tricolor (Tampala).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Amaranthaceae; Amaranthus.
OX NCBI_TaxID=29722;
RN [1]
RP SEQUENCE FROM N.A.
RA Iwamoto K., Fukuda H., Sugiyama M.;
RT "Elimination of POR Expression correlates with Red Leaf Formation in
RT Amaranthus tricolor.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB050119; BAB41186.1; -.
DR InterPro; IPR000343; GluTR.
DR Pfam; PF00745; GluTR; 1.
FT NON TER 1
FT NON TER 248
FT SEQUENCE 248 AA; 27814 MW; 99C31D03F7B4D4BB CRC64;

Query Match      34.9%; Score 910.5; DB 10; Length 248;
Best Local Similarity 72.5%; Pred. No. 7.6e-46;
Matches 177; Conservative 31; Mismatches 35; Indels 1; Gaps 1;

Qy 269 IGAGKMGKLVKHLVAKGCKVYVVRVNRVDAIREEMKDIEIVYRPLSDMYQAAA 328
Db 1 IGAGKMGKLVKHLVAKGCKVYVVRVNRVDAIREEMKDIEIVYRPLSDMYQAAA 328

Qy 329 VFTSTASSETSLFAKEHAELPPVSDTMGGVRLFVDISVPRNVSCVSEVGAARVYVND 388
Db 61 VFTSTASESLLPKENVQALPPVSGVGRFRFIDISVPRNVSGCISELENKAVYVND 120

Qy 389 LKEVVEANKEDRLRKAMEAQTIITIELRFEAWRDSLETVPYTKLRSYADRIRASELEK 448
Db 121 LKEVVAANKEDRLRKAMEAQAIITEIKFEAWRDSLETVPYTKLRSYADRIRASELEK 180

Qy 449 CLKQVGEDALTKMRRRAIEELSTGVNKLHGLQHLRCDGSDSRTLDETLENMHALNRM 508
Db 181 CLKSKMNP-LSKKQCAVDDLSGLNKLHGLQHLRCDGSDSRTLDETLENMHALNRM 239
```

```

QY 509 FSLD 512
DB 240 FSLE 243

RESULT 8
QBDI53
ID QBDI53 PRELIMINARY; PRT; 426 AA.
AC QBDI53;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Transfer RNA-Gln reductase.
GN HENA OR TL11738.
OS Synecococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BP-1;
RX MEDLINE=2225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriyuchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1."
RL DNA Res. 9:123-130(2002).
DR EMBL; AP005375; BAC09290.1; -.
KW Complete proteome.
SQ SEQUENCE 426 AA; 47596 MW; D84CESA1D2AA777E CRC64;

Query Match 34.3%; Score 893.5; DB 16; Length 426;
Best Local Similarity 44.2%; Pred. No. 1.6e-44;
Matches 193; Conservative 87; Mismatches 134; Indels 23; Gaps 5;

QY 82 IAVIGLSVHTAPVEMREKLAVALAEALPRAIOELTSLNHIIEAAVLSTCNRMIEIVVALSW 141
DB 3 IAVIGLSVHTAPVEMREKLAVALAEALPRAIOELTSLNHIIEAAVLSTCNRMIEIVVALSW 141
QY 142 NRGIREVVDWMSKSGIPASELREHLFILRSSDATRHLFEVSAGLSLVLGEGOILAQVK 201
DB 63 EVGVREHVQFLSEWSHIPQLRPYLFILLHQDAVMHLMRVASGLDSLIVIGEGILSQVK 122
QY 202 QVRSQNSGGLGNIDRMFKDAITAGKRVSETNISGAVSVSSAAVELAMLPKSEA 261
DB 123 RCHOLGOQYKSGIPILNRVFTGATAAGKRVETSTIGTGAVSISAAVELADRLQLQ- 181
QY 262 LSARMLLIGAGKMGKLVIKHLVAGCKVWVNRVSDAIREEMKDIEIVRPLSDMY 321
DB 182 -NCRIAVGAGKMSLVQVHLIARGVKIRIINSLERAQIAQOQFEVRFELFTMTDLL 240
QY 322 QAAAEADVFTSTASLSLFAKEHAELPPVSDTMGGVRL-----FVDISVPRNSACV 375
DB 241 PIVAAADLVFTSTAATEPLDR-----DNLGVLVGDRLAIDISVPRNVHANV 290
QY 376 SEVGAARYVVDLKEVVEANKEDRLRKAMEAQIITEELRRFEAWRDSLETPTIKKL 435
DB 291 TELGTVOLFNVDLQAVVAQNEARRQAQAEGLIEELETFLAWHALETVPRIISLR 350
QY 436 SYADIRIASELEKCLQKVGEDALTKMRRATEELSTGIVNKLHGLPQLHRCDSRSL 495
DB 351 QKMEAIRTOLEKALSRLGSE-FADKQGVIEAMTRTIINKILHDPVTQLOQ-----SORDL 405
QY 496 DETLENHMLNRMSLD 512
DB 406 ESRQAMQTLQDLFNL 422

RESULT 9
QBDFF8
ID QBDFF8 PRELIMINARY; PRT; 419 AA.
AC QBDFF8;

```

```

DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Glutamyl-tRNA reductase.
GN VW10254.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.S.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6."
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016797; AA008790.1; -.
KW Complete proteome.
SQ SEQUENCE 419 AA; 46057 MW; 834113C720656CBD CRC64;

Query Match 24.8%; Score 646; DB 16; Length 419;
Best Local Similarity 36.6%; Pred. No. 5e-30;
Matches 154; Conservative 81; Mismatches 144; Indels 42; Gaps 9;

QY 81 TIAVIGLSVHTAPVEMREKLAVALAEALPRAIOELTSLNHIIEAAVLSTCNRMIEIVVALS 140
DB 2 SLLAIGINHHTASVDLREKVAFGDPDKLANALQQLSQHEAVNGSVILSTCNRTVYCDVKS 61
QY 141 WNRGIREVVDWMSKSGIPASELREHLFILRSSDATRHLFEVSAGLSLVLGEGOILAQV 200
DB 62 GARS--KVIDWLSOFHQIGLEELAPSLYVVEEQAAIRHLMRVSCGLDSLVLGEPQLGQV 119
QY 201 KQVRSQNSGGLGNIDRMFKDAITAGKRVSETNISGAVSVSSAAVELA---LMLKP 257
DB 120 KQAVSDREQQAVDASLEKLFQKTFVSAKRVRTETDIGNAVSVAYAACTLAKHIFESLE 179
QY 258 KSEALSARMLLIGAKMGKLVIKHLVAGCKVWVNRVSDAIREEMKDIEIVRPL 317
DB 180 KS-----TVLLVAGAGETIELVAKHLHANGCKSMIVANRTREALTALAQF-DAQVI--SL 231
QY 318 SDMYQAAAEADVFTSTAS-----ETSLFAKEHAELPPVSDTMGGVRLFVDLSV 367
DB 232 QEINHLAKADIVISSTASPLPIGKGMVETALKQRHQP-----MLLVDIAV 279
QY 368 PRNVSACVSEGAARYVVDLKEVVEANKEDRLRKAMEAQIITEELRRFEAWRDSLET 427
DB 280 PRDVEAQVGLNDAYLTVDDLSQIIDSNIQKRVKVEAIQAEIQAIEAESAAFWMLRSLOA 339
QY 428 VPTIKKLRSYADRIASELEKCLQKVGEDALTKMRRATEELSTGIVNKLHGLPQLH 485
DB 340 VDSIREYROSANEIREDLKSLALLSLRSGSDP-----EKVLRLELSNRLTNKLIHAPTRAL 394
QY 486 R 486
DB 395 Q 395

RESULT 10
Q8GCB0
ID Q8GCB0 PRELIMINARY; PRT; 446 AA.
AC Q8GCB0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Glutamyl-tRNA reductase.
GN HEMA.
OS Bacillus megaterium.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1404;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM509;
RA Raux E., Leech H.K., Beck R., Schubert H.L., Santander P.J.,

```


RA Roesner C.A., Scott A.I., Thernes C., Rambach A., Martens J.H.,
 RA Barg H., Jahn D., Warren M.J.;
 RT "Identification and functional analysis of enzymes required for
 RT precorrin-2 dehydrogenation and metal ion insertion in the
 RT biosynthesis of sirohaem and cobalamin in *Bacillus megaterium*.";
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ508220; CAD48144.1;
 SQ SEQUENCE 446 AA; 49838 MW; 8A86F6782235A02 CRC64;

Query Match 23.7%; Score 618.5; DB 2; Length 446;
 Best Local Similarity 33.6%; Pred. No. 2.2e-28;
 Matches 148; Conservative 101; Mismatches 177; Indels 15; Gaps 7;

Qy 82 IAVIGLSVHTAPVEMREKLAVALPRAIOELTSLNHIEAAVLSTCNRMIEIYVALSW 141
 Db 3 IAVIGLSVHTAPVEMREKLAVALPRAIOELTSLNHIEAAVLSTCNRMIEIYVALSW 141
 Qy 142 NRGIREVVDWMSKSGIPASELREHLFILRSSDATRHLFEVSAGLDSLVLGEGQILAQVK 201
 Db 63 HTGRYYVKAFLAEWFGIDKEEPSYLTIVENDGAIEHLRYVACGLDSVIGETQILQVR 122
 Qy 202 QVVRSGQSGGLGKNDIMFKDAITAGKRVRSSTNISSGAVSVSSAAVELALMLKPKSEA 261
 Db 123 SSFLAQBEETIGTVFNOLFQOAVTLAKKAHETHEIGANAVSVSAAVELA--KKIFGDL 180
 Qy 262 LSARMLLIGAGKMGKLVKHLVAKGCKVWVNRVSVERVDAIREEMKDIEIYVRPLSD 321
 Db 181 SSKVLLIGAGKMGKLVKHLVAKGCKVWVNRVSVERVDAIREEMKDIEIYVRPLSD 321
 Qy 322 QAAEADVVFTSTASSETSLFAKEHAELPPVSDTMGGVRLF-VDISVPRNVSSACVSEVGA 380
 Db 238 HALSEADILISTGANDYVVTK--QMSAEARTKGRPLFWVDIAVRDLDPDELLET 294
 Qy 381 ARVYVVDLKEVVEANKEDRLKAMAEOTITIELRRFEARWDSLETPTIKLRSYADR 440
 Db 295 VFLYDIDLNGVNESLOKRADEIEIMLEAEIVAFKSWLGTGLGVVVISALRQKALT 354
 Qy 441 IRASELEKLOKVGEDALTKMRRRAIEELSTGIVNKLHGLPQLHRCDSRSLDETLE 500
 Db 355 IQAETMKSIDRLPP--LSERKRVNKHGTSIINQLLRDPHPAK--ELAGEKHAESLE 411
 Qy 501 NNHALNRMFSLDMKAEIIEQK 521
 Db 412 ---LPMKIFNIEQVALQKEK 429

RESULT 11
 QBEAR2
 ID Q8EAR2 PRELIMINARY; PRT; 416 AA.
 AC Q8EAR2;
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Glutamy1-tRNA reductase.
 GN HEMA OR SO3834.
 OS Shewanella oneidensis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
 OC Alteromonadaceae; Shewanella.
 OX NCBI_TaxID=70863;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MR-1;
 RX MEDLINE=2297686; PubMed=12368813;
 RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
 RA Read T.D., Eisen J.A., Seshadri R., Ward N., Meche B., Clayton R.A.,
 RA Meyer R.T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
 RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
 RA Vamathevan J., Weidman J., Imprim M., Lee K., Berry K., Lee C.,
 RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
 RA Feldblyum T.V., Smith H.O., Venter J.C., Nealeon K.H., Fraser C.M.;
 RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
 Shewanella oneidensis.";

RL Nat. Biotechnol. 20:1118-1123 (2002).
 DR EMBL: AE015816; AAN56811.1;
 DR TIGR: SO3834;
 KW Complete proteome.
 SQ SEQUENCE 416 AA; 45686 MW; 336289F059E50C57 CRC64;

Query Match 22.9%; Score 597; DB 16; Length 416;
 Best Local Similarity 35.0%; Pred. No. 3.7e-27;
 Matches 142; Conservative 86; Mismatches 162; Indels 16; Gaps 6;

Qy 81 TIAVIGLSVHTAPVEMREKLAVALPRAIOELTSLNHIEAAVLSTCNRMIEIYVALSW 140
 Db 2 SLVAIGLNHKTATVDLREKVAFPDKTHDANKSLASRSGEAVIVSTCNRTELYC----- 57
 Qy 141 WNRGIR-EVVDWMSKSGIPASELREHLFILRSSDATRHLFEVSAGLDSLVLGEGQILAQ 199
 Db 58 -NNGDETDIEWLEBYHGLEHQDVAPCLYNYHQGEAVRHLMRVASGLDSLILGEPOILQ 116
 Qy 200 QVVRSGQSGGLGKNDIMFKDAITAGKRVRSSTNISSGAVSVSSAAVELALMLKPKS 259
 Db 117 VKQAFVKAKEAGTVALTIDRLFQNTFSVAKKVRTETETIGAAAASVFAAASMA--KHIFS 174
 Qy 260 EALSARMLLIGAGKMGKLVKHLVAKGCKVWVNRVSVERVDAIREEMKDIEIYVRPLSD 319
 Db 175 SLSTTKVLLIGAGSTIELVAKHLKONGVAVSVNVRTILERAQSMCEEFNATAITLAIQIPD 234
 Qy 320 MYQAAEADVVFTSTASSETSLFAKEHAELPPVSDTMGGVRLFVDISVPRNVSSACVSEVG 379
 Db 235 F--LPRADIVISSTASPLPILGKMGVE--KALKQRRHQPMLLVDIAVRDIEPEVADLD 289
 Qy 380 AARVYVVDLKEVVEANKEDRLKAMAEOTITIELRRFEARWDSLETPTIKLRSYADR 439
 Db 290 DAFLYTVDDLHSIIEQNKASKEAAEQABLITEQSHLFMEWVRSLSDSIREYRSQSM 349
 Qy 440 RIRASELEKLOKVGEDALTKMRRRAIEELSTGIVNKLHGLPQLH 485
 Db 350 AIKDELVERALKLAQSGDTEQV---LVELANRLTNRLIHAFTQAL 392

RESULT 12
 Q8EPM7
 ID Q8EPM7 PRELIMINARY; PRT; 457 AA.
 AC Q8EPM7;
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Glutamy1-tRNA reductase (EC 1.2.1.1-).
 GN HEMA OR OB2070.
 OS Oceanobacillus iheyensis.
 OC Bacteria; Firmicutes; Bacillales; Oceanobacillus.
 OX NCBI_TaxID=182710;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HTE831 / DSM 14371 / JCM 11309;
 RX MEDLINE=22220767; PubMed=12235376;
 RA Takami H., Takaki Y., Uchiyama I.;
 RT "Genome sequence of *Oceanobacillus iheyensis* isolated from the Iheya
 RT Ridge and its unexpected adaptive capabilities to extreme
 RT environments.";
 RL Nucleic Acids Res. 30:3927-3935 (2002).
 DR EMBL: AP004599; BAC14026.1;
 KW Oxidoreductase; Complete proteome.
 SQ SEQUENCE 457 AA; 51544 MW; 2D936FCD7D9DFF4A CRC64;

Query Match 22.6%; Score 588; DB 16; Length 457;
 Best Local Similarity 32.7%; Pred. No. 1.4e-26;
 Matches 147; Conservative 90; Mismatches 195; Indels 18; Gaps 8;

Qy 82 IAVIGLSVHTAPVEMREKLAVALPRAIOELTSLNHIEAAVLSTCNRMIEIYVALSW 141
 Db 3 ILKGVFNKTPVDIREKFTFSDSLQDAMVELKNQKSILEDVIISTCNRTIYAVVDQL 62
 Qy 142 NRGIREVVDWMSKSGIPASELREHLFILRSSDATRHLFEVSAGLDSLVLGEGQILAQVK 201

Db 63 HTGRYIKQPLSNWFGIEKEEFSTYLRITEDDCAMEHLPRVSTGLDSWVLGETQILQVYK 122
QY 202 QVVRSGNSGGLKNIDRMFKDAITAGKRVRSSTNISSGAVSVSSAAVELALMKLPKSEA 261
Db 123 QAFNLSQGVNTTITFNEFLFKQAITFGKRAHKETAJGEHAVSISYAABELA--KKIFGDL 180
QY 262 LSARMLLIGAGKMGKLVKHLVAKGCKVNVNRSVERVDAIREEMK-DIEIVYRPLSDM 320
Db 181 QEGHVAILGAGKMGKLAANKIQSGGATKIVNRTLENANEMAEKNADVE-----SIDQL 236
QY 321 YQAAAEADVFTSTASSETSLFAKEHAEALPPVSDTGGVRLFVDISVPRNVSAACVSEVGA 380
Db 237 PVILQADILISSTGADSVIVTKEMMEKVQ--KORKGRALFLVDIAVPRDMDPAISELEN 294
QY 381 ARVYVDDILKEVVEANKEDRLRKAMEAQTIITEELRRFEAWRDSLETPTIKLRSYADR 440
Db 295 VFYLDIDNLOHIVDDNLEGRKQAKIELLIEEIVTFKEWLKTLGVIPVISALROKALT 354
QY 441 IRASELEKCIQKVGEDALTKMRRRAIEELSTGIWNKLLHGPLQHLRCDGSDSRTLDLE 500
Db 355 IQAETWQSIERTKIPN--LTDREKVLNKHTKSIIINQLLEPVTQAK----EPAGKDNASD 408
QY 501 NMHALNMFSLDMKEAIIQKIK-AKVEKT 529
Db 409 SLQLFNIFQI--EEBVKSELVHAKKNET 436

RESULT 13

Q8FY103 PRELIMINARY; PRT; 484 AA.
AC Q8FY103;
DT 01-MAR-2003 (TREMREL. 23, Created)
DT 01-MAR-2003 (TREMREL. 23, Last sequence update)
DE Glutamy1-trNA reductase (EC 1.2.1.1.).
GN HEMA OR C1668.
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Raetz D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.F., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli".
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR EMBL; AF016759; JN80133.1; --
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 484 AA; 53679 MW; 462CAB49487DC59 CRC64;

Query Match 21.9%; Score 570.5; DB 16; Length 484;
Best Local Similarity 33.7%; Pred. No. 1.6e-25;
Matches 141; Conservative 88; Mismatches 171; Indels 19; Gaps 8;

QY 81 TIATVGLSVHTAPVEMREKLAVALPRAIQELTSLNHIIEAAVLSTCNRMIEIYVALS 140
Db 68 TLLALGINHTKAPVSLRERSVSPDKLDQALDSLLAPQVGVVLTSCNRTLYLVEE 127
QY 141 WNRGIREVVDMKSGIPASELREHLFILRSSDATRHLFEVSAGLSDSLVLEGQILAQV 200
Db 128 RDQLQALIRWUDDYHNLNEDDLRSLYWHQNDVASHLMRVASGLSDSLVLEGQILGV 187
QY 201 KOVVRSGNSGGLKNIDRMFKDAITAGKRVRSSTNISSGAVSVSSAAVELALMKLPKSEA 260
Db 188 KKAFADSQGHMKASELERMFQKSFSAKRVATETDIDGASVSVFAACTLARQIF---E 244
QY 261 ALS-ARMLLIGAGKMGKLVKHLVAKGCKVNVNRSVERVDAIREEMKDIEIVYRPLSD 319

Db 245 SLSTVTVLVGAGETIELVARHLREHKVQKMIANTRTERRAQILADEV-GAEVI--ALSD 301
QY 320 MYQAAAEADVFTSTASSETSLFAKEHAEALPPVSDTGGVRLFVDISVPRNVSAACVSEVG 379
Db 302 IDERLEADIIISSTASPLPIIGKGMVER--ALKSRNQPMILLIVIAVPRDVEPEVGKLA 359
QY 380 ARVYVDDILKEVVEANKEDRLRKAMEAQTIITEELRRFEAWRDSLETPTIKLRSYAD 439
Db 360 NAYLSVDDQLGIIISHINLAQRKAAVAEETIIVAQEASEFMAMURAOSASSETIRDSQAE 419
QY 440 RIRASELEKCIQKVGEDALTKMRRRAIEELSTGIWNKLLHGPL--LQHLRCDGSDSR 493
Db 420 QYRDELTAALAALEGGDAQT-----IMQDLAKWLTNRLIHAPTLSLQQAARDGDNER 473

RESULT 14

Q8CNY6 PRELIMINARY; PRT; 448 AA.
AC Q8CNY6;
DT 01-MAR-2003 (TREMREL. 23, Created)
DT 01-MAR-2003 (TREMREL. 23, Last sequence update)
DE Glutamy1-trNA reductase.
GN SE1347.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,
RA Chen Z., Wen Y.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016748; AA004946.1; --
KW Complete proteome.
SQ SEQUENCE 448 AA; 50151 MW; F89F44784F3256A2 CRC64;

Query Match 20.9%; Score 543.5; DB 16; Length 448;
Best Local Similarity 30.5%; Pred. No. 5.6e-24;
Matches 137; Conservative 103; Mismatches 182; Indels 27; Gaps 9;

QY 85 IGLSVHTAPVEMREKLAVALPRAIQELTSLNHIIEAAVLSTCNRMIEIYVVALSWNRG 144
Db 6 ISINHRTADVLRQSVAFRDDALRLAHEDLYETKAILLENVILSTCNRTVEYAIVDQVHTG 65
QY 145 IREVVDMKSGIPASELREHLFILRSSDATRHLFEVSAGLSDSLVLEGQILAQVQV 204
Db 66 RYIQRFLARSFGFEVDIDIKMSQVKGVDGDAVEHLLRVTSGLSDSLVLEGQILQNRDAP 125
QY 205 RSGQNSGGLKNIDRMFKDAITAGKRVRSSTNISSGAVSVSSAAVELALMKLPKSELSA 264
Db 126 FLAQNTGTGTITFNHLFKQAITFAKKAHSETDIADNAVSVSYAAVELAKKVFGLK--SK 183
QY 265 RMLLIGAGKMGKLVKHLVAKGCKVNVNRSVERVDAIREEMKDIEIVYRPLSDMYQA 324
Db 184 HAVVLGAGEMELSLNLLGSGISNVTIVNRTLSKAKILAEKH---NVSYDSLALPSLL 240
QY 325 AEADVFTSTASSETSLFAKEHAEALPPVS--DTMGVRLFVDISVPRNVSAACVSEVGAAR 382
Db 241 ETTDVISSTSAEDYIITNSVTKTISERKLDLSL---VLIDIAVPRDIEPIDAITNIF 296
QY 383 VYVDDILKEVVEANKEDRLRKAMEAQTIITEELRRFEAWRDSLETPTIKLRSYAD 439
Db 297 NYDVDDLDLVLDANLRER---QLAAETIAGQIPEIDSHNENWVMGLVVPVIRALREKAM 353
QY 440 RIRASELEKCIQKVGEDALTKMRRRAIEELSTGIWNKLLHGPLQHLRCDGSDSRTLDL 499
Db 354 NIQETMESIDRKLPD--LSERERKVISKHTKSIINQMLKPIQAKELSTDKK----SN 407
QY 500 ENMHALNMFSLDMKEAIIQKIKAKVEK 528
Db 408 EKLELFQWIFDIEAE---DPRERAKLEK 432

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 12:53:42 ; Search time 47 Seconds
(without alignments)
1793.273 Million cell updates/sec

Title: US-10-018-902-4

Perfect score: 2606

Sequence: 1 MATTTTATTAATAAATAATKAP.....DMEKALIEQIKAKVETQXN 531

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2606	100.0	531	22 AAB31960	Amino acid sequenc
2	2272	87.2	480	22 AAB31962	Amino acid sequenc
3	2091	80.2	454	22 AAB31968	Amino acid sequenc
4	1783.5	68.4	536	22 AAB31964	Amino acid sequenc
5	1770.5	67.9	543	23 ABB31383	Herbicidally activ
6	1677	64.4	530	23 ABB30929	Herbicidally activ
7	1540	59.1	467	22 AAB31966	Amino acid sequenc
8	1334	51.2	524	23 ABB31959	Herbicidally activ
9	637.5	24.5	415	21 AAY75260	Neisseria meningit

10	632.5	24.3	415	21 AAY75259	Neisseria gonorrhoe
11	632.5	24.3	415	24 AAB78975	N. gonorrhoeae ami
12	628.5	24.1	415	21 AAY75261	Neisseria meningit
13	594.5	22.8	426	16 AAR74619	Xanthomonas L-glut
14	543.5	20.9	448	22 AAG82317	S. epidermidis ope
15	543.5	20.9	448	23 ABB38210	Staphylococcus epi
16	506	19.4	435	23 ABB48846	Listeria monocytog
17	483.5	18.6	463	22 AAG90216	C glutamicum prote
18	428.5	16.4	449	19 AAW98821	H. pylori GHPO 142
19	388	14.9	87	22 AAB31961	Amino acid sequenc
20	282	10.8	63	22 AAB31959	Amino acid sequenc
21	219.5	8.4	384	22 AAU67473	Propionibacterium
22	206.5	7.9	341	20 AAY35341	Chlamydia pneumonia
23	195.5	7.5	348	20 AAY37517	Protein involved i
24	145.5	5.6	1764	10 AAP91672	Primary amino acid
25	139.5	5.4	975	19 AAW72746	Drosophila kinesin
26	139.5	5.4	975	22 ABB63485	Drosophila melanog
27	139	5.3	1275	22 AAW78921	Human protein SEQ
28	139	5.3	1301	22 AAW79905	Human protein SEQ
29	139	5.3	2017	22 ABB66301	Novel human diagno
30	138	5.3	1725	21 AAB19800	Mouse laminin 2 ma
31	138	5.3	1725	21 AAB48451	Mouse laminin 8 po
32	138	5.3	1725	23 ABB81593	Mouse laminin 10 s
33	138	5.3	1786	21 AAB19799	Mouse laminin 2 be
34	138	5.3	1786	21 AAB48450	Mouse laminin 8 po
35	138	5.3	1786	23 ABB81592	Mouse laminin 10 s
36	136.5	5.2	1776	19 AAW50894	Mouse laminin Bl c
37	135	5.2	3338	22 ABB09919	Novel human diagno
38	133	5.1	2563	23 AAU10873	Penicillium citrin
39	132.5	5.1	871	23 ABB39134	Staphylococcus epi
40	132	5.1	727	21 AAG47308	Arabidopsis thalia
41	129.5	5.0	1123	22 ABB12514	Novel human diagno
42	129.5	5.0	1123	22 ABB14529	Novel human diagno
43	129	5.0	648	23 ABB73703	Candida albicans e
44	129	5.0	1177	22 AAB96721	Putative P. abyssi
45	128.5	4.9	2757	22 ABB09921	Novel human diagno

ALIGNMENTS

RESULT 1

AAB31960
ID AAB31960 standard; Protein; 531 AA.

AC AAB31960;

DT 15-MAY-2001 (first entry)

DE Amino acid sequence of a corn Glu-tRNA reductase enzyme.

KW Glu-tRNA reductase; aminolevulinic acid; 5-aminolevulinic acid; ALA;
tetrapyrrolic pigment; chlorophyll; transgenic plant.

OS Zea mays.

PN W0200109304-A2.

PD 08-FEB-2001.

PF 28-JUL-2000; 2000WO-US21008.

PR 30-JUL-1999; 99US-0146600.

PA (DUPO) DU FONT DE NEMOURS & CO E I.

PI Cahoon RE, Gutteridge S, Harvell LT, Rafalski JA, Tao Y, Weng Z;

XX WPI; 2001-159865/16.

DR N-PSDB; AAF54814.

XX New nucleic acids encoding an aminolevulinic acid biosynthetic enzyme
PT useful for producing large amounts of the encoded polypeptides used in


```

Db      361  TQELKRFKAWDSLETPTIKLSYADRTASLEKCLKIGDALTKKRRSIBELST 420
Qy      472  GIVNKLHGPIQLHRCDSRSTLDETLENHMRMPSLDMKXAIIEQKIKAKVEKTON 531
Db      421  GIVNKLHGPIQLHRCDSRSTLDETLENHMRMPSLDMKXAIIEQKIKAKVEKSON 480

RESULT 3
AAB31968
ID      AAB31968 standard; Protein; 454 AA.
XX
AC      AAB31968;
XX
DT      15-MAY-2001 (first entry)
XX
DE      Amino acid sequence of a wheat Glu-tRNA reductase enzyme.
XX
KW      Glu-tRNA reductase; aminolevulinic acid; 5-aminolevulinic acid; ALA;
KW      tetrapyrrolic pigment; chlorophyll; transgenic plant.
XX
OS      Triticum aestivum.
XX
FN      WO200109304-A2.
XX
PD      08-FEB-2001.
XX
PF      28-JUL-2000; 2000WO-US21008.
XX
PR      30-JUL-1999; 99US-0146600.
XX
PA      (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI      Cahoon RE, Gutteridge S, Harvell LT, Rafalski JA, Tao Y, Weng Z;
DR      WPI; 2001-159865/16.
DR      N-PSDB; AAF54822.
XX
PT      New nucleic acids encoding an aminolevulinic acid biosynthetic enzyme
PT      useful for producing large amounts of the encoded polypeptides used in
PT      screening compounds for potential herbicidal activity -
XX
PS      Disclosure; Page 67-69; 77pp; English.
XX
CC      The present sequence represents a Glu-tRNA reductase, which is an
CC      aminolevulinic acid biosynthetic enzyme. The enzyme converts
CC      Glu-tRNA-Glu to glutamate 1-semialdehyde (GSA) with the concomitant
CC      release of tRNA-Glu. GSA aminotransferase then converts GSA to
CC      5-aminolevulinic acid (ALA). ALA is used in the biosynthesis of
CC      tetrapyrrolic pigments such as chlorophyll. Nucleic acids encoding
CC      aminolevulinic acid biosynthetic enzymes are useful for isolating cDNAs
CC      and genes encoding homologous proteins from the same or other plant
CC      species, for creating transgenic plants in which the polypeptides are
CC      present at higher or lower levels than normal or in cell types or
CC      developmental stages in which they are not normally found, for
CC      overexpression in bacterial or yeast hosts to efficiently produce
CC      large amounts of the encoded polypeptides which could then be used for
CC      screening different compounds for potential herbicidal activity, and
CC      as hybridisation probes and amplification primers.
SQ      Sequence 454 AA;

Query Match      80.2%; Score 2091; DB 22; Length 454;
Best Local Similarity 91.4%; Pred. No. 3e-169;
Matches 415; Conservative 25; Mismatches 14; Indels 0; Gaps 0;

Qy      78  ERSTIAVIGLSVHTAPVEMREKLAVAEELMPRAIQELTSLNHIEAAVLSTCNRMELIYV 137
Db      1  EKSSIAVIGLSVHTAPVDMREKLAVAEELMPRAISELTSLNHIEAAVLSTCNRMELIYV 60
Qy      138  ALSWNRGIRVVDWMSKSGIPASELREHLFILRSSDATRHLFEVSAGLSVLGEGQIL 197
Db      61  ALSWNRGIRVVDWMSKSGIPASELREHLFMLRDSATRHLFEVSAGLSVLGEGQIL 120

```

```

Qy      198  AQVKQVVRSGNSGGLGKNIDRMFKDAITAGKRVRSSETNISGAVSVSSAAVELAMKLP 257
Db      121  AQVKQVVRNGNSGGLGKNIDRMFKDAITAGKRVRCETNISAGAVSVSSAAVELAMKLP 180
Qy      258  KSEALSARMELLIGAGKMGKLVIKHLVAKGCKKVVVNRVSVERVDAIREMKDIEIVYRPL 317
Db      181  KSECLSARMELLIGAGKMGKLVVRLIIANGCKKVVVNRVSVERVDAIREMKDIEIVYRPL 240
Qy      318  SDMYQAAAADVVFTSTASSETSLFAKEHAEALPPVSDTMGGVRLFDIVSPRNVSAVCSE 377
Db      241  TEMVEAAAADVVFTSTASSETSLFTEHAEALPPISLAVGGVRLFDIVSPRNVGACVSE 300
Qy      378  VGARVYVNVDDLKEVVEANKEDRLRKAEBAQTIIIEELRRFEAWRDSLETPTIKKLSY 437
Db      301  VEHARVYVNVDDLKEVVEANKEDRVRKAEBAQTIIITQELKRFKAWRDSLETPTIKKLSY 360
Qy      438  ADRTRASELEKCLKQVGEDALTKKRRRAIEELSTGIYNKLLHGPIQLHRCDSRSTLDE 497
Db      361  ADRTRASELEKCLKQIGEDNLNKKRRSIEELSTGIYNKLLHGPIQLHRCDSRSTLDE 420
Qy      498  TLENMHALNRMFSLDMKXAIIEQKIKAKVEKTON 531
Db      421  TLENMHALNRMFNLDTKAVLEQKIKAKVEKTON 454

RESULT 4
AAB31964
ID      AAB31964 standard; Protein; 536 AA.
XX
AC      AAB31964;
XX
DT      15-MAY-2001 (first entry)
XX
DE      Amino acid sequence of a soybean Glu-tRNA reductase enzyme.
XX
KW      Glu-tRNA reductase; aminolevulinic acid; 5-aminolevulinic acid; ALA;
KW      tetrapyrrolic pigment; chlorophyll; transgenic plant.
XX
OS      Glycine max.
XX
FN      WO200109304-A2.
XX
PD      08-FEB-2001.
XX
PF      28-JUL-2000; 2000WO-US21008.
XX
PR      30-JUL-1999; 99US-0146600.
XX
PA      (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI      Cahoon RE, Gutteridge S, Harvell LT, Rafalski JA, Tao Y, Weng Z;
DR      WPI; 2001-159865/16.
XX
PT      New nucleic acids encoding an aminolevulinic acid biosynthetic enzyme
PT      useful for producing large amounts of the encoded polypeptides used in
PT      screening compounds for potential herbicidal activity -
XX
PS      Claim 2; Page 55-57; 77pp; English.
XX
CC      The present sequence represents a Glu-tRNA reductase, which is an
CC      aminolevulinic acid biosynthetic enzyme. The enzyme converts
CC      Glu-tRNA-Glu to glutamate 1-semialdehyde (GSA) with the concomitant
CC      release of tRNA-Glu. GSA aminotransferase then converts GSA to
CC      5-aminolevulinic acid (ALA). ALA is used in the biosynthesis of
CC      tetrapyrrolic pigments such as chlorophyll. Nucleic acids encoding
CC      aminolevulinic acid biosynthetic enzymes are useful for isolating cDNAs
CC      and genes encoding homologous proteins from the same or other plant
CC      species, for creating transgenic plants in which the polypeptides are
CC      present at higher or lower levels than normal or in cell types or
CC      developmental stages in which they are not normally found, for
CC      overexpression in bacterial or yeast hosts to efficiently produce

```


PI Tettelin H, Venter JC;
 XX WPI; 2000-062150/05.
 DR N-PSDB; AA254022.
 XX
 PT Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics -
 XX
 PS Claim 2; Page 996; 1453pp; English.
 XX
 CC AA253015 to AA254536, AA254577 to AA254615, and AA254615 to AA254615
 CC represent novel Neisseria meningitis and N. gonorrhoea polynucleotides
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA254615 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of
 CC the invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the
 CC manufacture of medicaments for treating or preventing infection due to
 CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
 CC presence of Neisseria bacteria, or to raise antibodies. They may also
 CC be used to screen for agonists or antagonists, which may themselves
 CC have use as antibacterial agents. The polynucleotides of the invention
 CC may also be used in gene therapy protocols.
 XX
 SQ Sequence 415 AA;
 Query Match 24.5%; Score 637.5; DB 21; Length 415;
 Best Local Similarity 34.3%; Pred. No. 1.5e-45;
 Matches 149; Conservative 96; Mismatches 163; Indels 27; Gaps 9;
 QY 82 IAVIGLSVHTAPVEMREKLAVALAEELPRAIQELTSLNHIIEAAVLSTCNRMIEIYVVALSW 141
 Db 3 LTAVGLNHOTAPLSIREKLAFAAAALPRAVNLARSNAATEAVILSTCNRTLYCVGDS- 61
 QY 142 NRGIREVVDWMSKSGIPASELREHLFILRSSDATRHLFEVSAGLDSLVLGEGQILAQVK 201
 Db 62 ----EELIRWLADYHSLPIEIRPYLYALDMQETVRHAFRVACGLDSVLGEPQLGQIK 117
 QY 202 QVVRSGQSGGLGNIDRMFKDAITAGKRVSRSETNISSGAVSVSSAAVELALMKLPKSEA 261
 Db 118 DAVRAQEQESMGKLNALFQKTFSAKEVTRDTAVGENSVMSASVLAELQIFPDIGD 177
 QY 262 LSARMLLIGAKMGKLVKHLVAKGCKKVVVNRVSVERVDAIREMKD-IBIVYRP--LS 318
 Db 178 LN--VLFIGAGEMIELVATYFAAKSPRLMTVANRTLARA----QELCDKLGVAEPCLLS 231
 QY 319 DMYQAAAEADVFTSTASLSLFAKEHAELPPVSDTMGGVRLP-VDISVPRNVSAVCSE 377
 Db 232 DLPAILHDYDVVSVSTASQLPIVGKMYERALKQKQSM---PLFMLDLAVPRDIEAEVGD 288
 QY 378 VGAARVYVNDLKEVVEANKEDRLRKAMEAQTITTEELRRFEAWRDSLETPTIKLRSY 437
 Db 289 LNDAYLYTVDDWNVIVSGKARQAKAAATAATLVSEKVAEFVROOQGRQSVPLIKALRDE 348
 QY 438 ADRTASELEKCLKQVGEDALTQKORRAIEBELSTGIVKNLHGLPQHLRCGDSRSLDE 497
 Db 349 GEKARKQVLENAMKQAKGATAEEV---LERLSVQLTNKLHLSPTQTLNKAEGEDKOL-- 403
 QY 498 TLENKHALNRNFSLD 512
 Db 404 ----VHAQAQIYHLD 414
 RESULT 10
 AA254536
 ID AA254536 standard; Protein; 415 AA.
 XX
 AC AA254536;
 XX
 DT 21-MAR-2000 (first entry)
 XX
 DE Neisseria gonorrhoea ORF 622 protein sequence SEQ ID NO:1992.
 XX

Neisseria meningitidis; Neisseria gonorrhoea; antigen; vaccine;
 antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
 antibacterial; gene therapy.
 Neisseria gonorrhoea.
 WO9957280-A2.
 11-NOV-1999.
 30-APR-1999; 99WO-US09345.
 01-MAY-1998; 98US-0083758
 31-JUL-1998; 98US-0094869.
 02-SEP-1998; 98US-0098994.
 02-SEP-1998; 98US-0099062.
 09-OCT-1998; 98US-0103749.
 09-OCT-1998; 98US-0103794.
 09-OCT-1998; 98US-0103796.
 25-FEB-1999; 99US-0121528.
 (CHIR) CHIRON CORP.
 (GENO-) INST GENOMIC RES.
 Frazer C, Galeotti C, Grandi G, Hickey E, Massignani V, Mora M;
 Petersen J, Pizzi M, Rappuoli R, Ratti G, Scialato E, Scarselli M;
 Tettelin H, Venter JC;
 WPI; 2000-062150/05.
 N-PSDB; AA254021.
 Novel Neisserial polypeptides predicted to be useful antigens for
 vaccines and diagnostics -
 Claim 2; Page 995; 1453pp; English.
 AA253015 to AA254536, AA254577 to AA254615, and AA254615 to AA254615
 represent novel Neisseria meningitis and N. gonorrhoea polynucleotides
 and polypeptides. AA254537 to AA254576 and AA254616 to AA254615 represent
 PCR primers used in the exemplification of the present invention. The
 polypeptides, the polynucleotides, antibodies and compositions of
 the invention can be used as vaccines, as diagnostic reagents, and as
 immunogenic compositions. The polypeptides can be used in the
 manufacture of medicaments for treating or preventing infection due to
 Neisserial bacteria (e.g. meningitis and septicemia), to detect the
 presence of Neisseria bacteria, or to raise antibodies. They may also
 be used to screen for agonists or antagonists, which may themselves
 have use as antibacterial agents. The polynucleotides of the invention
 may also be used in gene therapy protocols.

Query Match 24.3%; Score 632.5; DB 21; Length 415;
 Best Local Similarity 34.0%; Pred. No. 3.9e-45;
 Matches 148; Conservative 96; Mismatches 164; Indels 27; Gaps 9;
 QY 82 IAVIGLSVHTAPVEMREKLAVALAEELPRAIQELTSLNHIIEAAVLSTCNRMIEIYVVALSW 141
 Db 3 LTAVGLNHOTAPLSIREKLAFAAAALPRAVNLARSNAATEAVILSTCNRTLYCVGDS- 61
 QY 142 NRGIREVVDWMSKSGIPASELREHLFILRSSDATRHLFEVSAGLDSLVLGEGQILAQVK 201
 Db 62 ----EELIRWLADYHSLPIEIRPYLYTLDQMOTVRHAFRVACGLDSVLGEPQLGQIK 117
 QY 202 QVVRSGQSGGLGNIDRMFKDAITAGKRVSRSETNISSGAVSVSSAAVELALMKLPKSEA 261
 Db 118 DAVRAQEQESMGKLNALFQKTFSAKEVTRDTAVGENSVMSASVLAELQIFPDIGD 177
 QY 262 LSARMLLIGAKMGKLVKHLVAKGCKKVVVNRVSVERVDAIREMKD-IBIVYRP--LS 318
 Db 178 LN--VLFIGAGEMIELVATYFAAKNPRLMTVANRTLARA----QELCDKLGVAEPCLLS 231
 QY 319 DMYQAAAEADVFTSTASLSLFAKEHAELPPVSDTMGGVRLP-VDISVPRNVSAVCSE 377

Db 232 DLPAILHDYDVVVSTASQLPVIGKGMVERALKQKQSM---PLFMLDLAVPRDIEAEVGD 288
 Qy 378 VGAARVYNVDLKEYVEANKEDRLRKAMEAQTIITEELRRFEAWRDSLETPTIKLRSY 437
 Db 289 LNDAYLYTVDDMVNIVQSGKEARQKAAAAAETLVSEKVAEFVROQOGRQSVPLIKALRDE 348
 Qy 438 ADRIASELEKCLKQVGEDALTCKMRRRAIBELSTGIVNKLHLGPHLQHCDSRSTLDE 497
 Db 349 GEKARKQVLENAMKQAKGATAEEV---LERLSVLTNKLHLSPTQTLNKAGEEDKOL-- 403
 Qy 498 TLENMHALNRMFSLD 512
 Db 404 ----VHAVAQIYHLD 414

RESULT 11

ABP78975
 ID ABP78975 standard; Protein; 415 AA.

AC ABP78975;
 XX 07-MAR-2003 (first entry)
 DT N. gonorrhoeae amino acid sequence SEQ ID 4480.
 DE Antibacterial; infection; vaccine; gene therapy.
 KW Neisseria gonorrhoeae.
 OS Neisseria gonorrhoeae.
 PN WO200279243-A2.
 XX 10-OCT-2002.

PF 12-FEB-2002; 2002MO-IB02069.

PR 12-FEB-2001; 2001GB-0003424.

PA (CHIR-) CHIRON SPA.

PI Fontana MR, Pizza M, Massignani V, Monaci E;

XX WPI; 2003-058415/05.

DR N-PSDB; ABZ39945.

PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
 medicament for treating or preventing N. gonorrhoeae infection -

XX Disclosure; Page 516; 815pp; English.

XX The present invention relates to proteins from Neisseria gonorrhoeae.
 CC Also disclosed are the nucleic acid molecules encoding the proteins and
 CC antibodies that specifically bind to the proteins. The composition
 CC comprising the protein, nucleic acid or antibody is useful for the
 CC manufacture of a medicament for treating or preventing N. gonorrhoeae
 CC infection, this may be in the form of a vaccine or gene therapy.
 CC Sequences given in records ABP76736-ABP81046 represent nucleic acid
 CC molecules of the invention.

XX Sequence 415 AA;

Query Match 24.3%; Score 632.5; DB 24; Length 415;
 Best Local Similarity 34.0%; Pred. No. 3.9e-45;
 Matches 148; Conservative 96; Mismatches 164; Indels 27; Gaps 9;

Qy 82 IAVIGLSVHTAPVEMREKLAVALAEELWPRATQELTSLNHIIEAAVLSTCNRMETIYVVALSW 141

Db 3 LTAVGLNHQTAPLSIREKLAFAAAALPEAVRNLRASNAATEAVILSTCNRTLEYCVGDS- 61

Qy 142 NRGIREVVDNMKSGIPASELREHLFILRSSDATRHLFEVSAGLSLVLGEGQILAOVK 201

Db 62 ----EIRLWADYHSUFIEIRPLYLTLDQETVRHAFRVACGLDSMVLGEPQILGQIK 117

Qy 202 QVTRSGNSGLGNKIDRMFKDAITAGKVRSETNISGAVSVSSAAVELALMKLPKSEA 261
 Db 118 DAVRAAQEQESMGAKNALFQKTESVAKEVRTDTAVGENSVSMASVVKLAEQIFPDIGD 177
 Qy 262 LSABMLLIGACKMCKLVKHLVAKGCKKVVVVRVSVERVDAIREMKD-IEIVYRP--LS 318
 Db 178 LN---VLFIGAMELVELVATYFAAKNPRMTVANRTLARA----QELCDKGVNAEPCLLS 231
 Qy 319 DMYQAAAEADVFTSTASSETSLFAKEHAEALPPVSDTMGGVRLF-VDISVPRNVSYACVSE 377
 Db 232 DLPAILHDYDVVVSTASQLPVIGKGMVERALKQKQSM---PLFMLDLAVPRDIEAEVGD 288
 Qy 378 VGAARVYNVDLKEYVEANKEDRLRKAMEAQTIITEELRRFEAWRDSLETPTIKLRSY 437
 Db 289 LNDAYLYTVDDMVNIVQSGKEARQKAAAAAETLVSEKVAEFVROQOGRQSVPLIKALRDE 348
 Qy 438 ADRIASELEKCLKQVGEDALTCKMRRRAIBELSTGIVNKLHLGPHLQHCDSRSTLDE 497
 Db 349 GEKARKQVLENAMKQAKGATAEEV---LERLSVLTNKLHLSPTQTLNKAGEEDKOL-- 403
 Qy 498 TLENMHALNRMFSLD 512
 Db 404 ----VHAVAQIYHLD 414

RESULT 12

AA75261
 ID AA75261 standard; Protein; 415 AA.

AC AA75261;

XX 21-MAR-2000 (first entry)

DE Neisseria meningitidis ORF 622 protein sequence SEQ ID NO:1996.

KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
 KW antibacterial; gene therapy.

OS Neisseria meningitidis.

XX WO9957280-A2.

XX 11-NOV-1999.

XX 30-APR-1999; 99WO-US09346.

XX 01-MAY-1998; 98US-0083758.

XX 31-JUL-1998; 98US-0094869.

XX 02-SEP-1998; 98US-0098994.

XX 02-SEP-1998; 98US-0099062.

XX 09-OCT-1998; 98US-0103749.

XX 09-OCT-1998; 98US-0103794.

XX 09-OCT-1998; 98US-0103796.

XX 25-FEB-1999; 99US-0121528.

XX (CHIR) CHIRON CORP.

PA (GENO-) INST GENOMIC RES.

XX Fraser C, Galeotti C, Grandi G, Hickey E, Massignani V, Mora M;

PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;

XX Tettelin H, Venter JC;

DR WPI; 2000-062150/05.

DR N-PSDB; AAZ54023.

XX Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics -

XX Claim 2; Page 997; 1453pp; English.

CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AA74253 to AA75941
 CC represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides


```

XX PN W0200134809-A2.
XX PD 17-MAY-2001.
XX PF 09-NOV-2000; 2000WO-US30782.
XX PR 09-NOV-1999; 99US-0164258.
XX PA (GLAX ) GLAXO GROUP LTD.
XX PI Kimmerly WJ;
XX DR WPI; 2001-316495/33.
XX DR N-PSDB; AAH53167.
XX PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
XX PT useful for vaccinating against infections, e.g. endocarditis -
XX PS Claim 18; Page 478; 2188pp; English.
XX CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
XX CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
XX CC (I) and (II) can have antibacterial activity and therefore can be used
XX CC in vaccination. The nucleic acids (I) may be used to produce the
XX CC S. epidermidis polypeptides (II) via the production of vectors
XX CC containing them which are used to produce hosts cells which express the
XX CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
XX CC used to vaccinate subjects and to raise antibodies against the bacteria.
XX CC The polypeptides may also be used to assay for other inhibitors of their
XX CC activity and therefore identify compounds that may be used for the
XX CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
XX CC AAH55030 represent specifically claimed S. epidermidis genomic DNA
XX CC polynucleotide sequences from the present invention. AAH55031 to
XX CC AAH55098 represent oligonucleotide sequences and primers which are used
XX CC in the exemplification of the present invention.
XX CC N.B. The present invention specifically claims all the polynucleotide
XX CC sequences given in the sequence listing of the present specification,
XX CC however the sequence listing only goes up to SEQ ID NO:4454 so even
XX CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
XX CC no sequences are present for SEQ ID NO:4455 to 4464.
XX SQ Sequence 448 AA;

Query Match 20.9%; Score 543.5; DB 22; Length 448;
Best Local Similarity 30.5%; Pred. No. 1.7e-37;
Matches 137; Conservative 103; Mismatches 182; Indels 27; Gaps 9;

QY 85 IGLSVHTAPVEMREKLAELPRAIQELTSLNHIEEAAVLSTCNRMIEIYVVALSWNRG 144
DB 6 ISINHRTADVTLEQVAFRDALRLAHEDLYETKAILLENVILSTCNRTVEIVAVDQVHTG 65

QY 145 IREVDWMKSKGIPASELREHLFILRSSDATRHLFEVSGAGLSLVGEGQILAQVKQV 204
DB 66 RYIQRFLARSGFGEVDVDDIKDMSEVKVGDVAVEHLRVTSGLSIVLGETQILQOMRDAF 125

QY 205 RSGQSGGLGNIDRMFKDAITAGKRVSRSETNISGSAVSVSSAAVELALMKLPKSEALSA 264
DB 126 FLAQTGTGTGIFNHLFKQAIFAKKAHSETDIADNAVSVSYAAVELAKKVFGLK--SK 183

QY 265 RMLLIGAGKMGKLVKHLVAKGCKVNVNRSVERVDAIREEMKDIEIVRPLSDMYQAA 324
DB 184 HAVVIGAGEMGELSLLNLGSGISNVTIVNRTLSKAILAEKH--NVSYDSLALPSLL 240

QY 325 AADVFTTASSETSLFAKEHAELPPVS--DTWGGVRLFVDISVPRNVSAACVSEVGAR 382
DB 241 ETTDIVISSTSAEDYIITNSVMVKTISETRKLDSL--VLIDIAVPRDIEFGDAITNIF 296

QY 383 VYNVDDLKEVVEANKEDRLKAMEAQT--ITEELRRFEAWRDSLETVPITIKLRSYAD 439
DB 297 NYVDLDLKDLYDANLRER--QLAETIAGQIPEDISHNEWNVLGVVPIVIRALREKAM 353
QY 440 RIRASELEKCLQKVGEDALTKOMRRAIEELSTGIVNKLHLGPHQLRCDGSDSRSLTDETL 499

```

```

DB 354 NIQAETWESIDRKLPD--LSERERKVISKTKTSIINQMLKDPKOAKELSTDKK-----SN 407
QY 500 ENMHALNRMFSLDMKEKAIIEOKIKAKVEK 528
DB 408 EKLELFQNIQFIDIEAE----DPREKAKLEK 432

RESULT 15
ABP38210
ID ABP38210 standard; Protein; 448 AA.
XX AC ABP38210;
XX DT 24-JUL-2002 (first entry)
XX DB Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3055.
XX KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
XX KW antibacterial; gene therapy.
XX OS Staphylococcus epidermidis.
XX PN US6380370-B1.
XX PD 30-APR-2002.
XX PF 13-AUG-1998; 98US-0134001.
XX PR 14-AUG-1997; 97US-055779P.
XX PR 08-NOV-1997; 97US-064964P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Doucette-Stamm LA, Bush D;
XX DR WPI; 2002-381255/41.
XX DR N-PSDB; ABN90755.
XX PT Novel isolated nucleic acid encoding a Staphylococcus epidermidis
XX PT polypeptide, useful for diagnosing and treating bacterial infections -
XX PS Disclosure; SEQ ID 3055; 267pp; English.
XX CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
XX CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
XX CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
XX CC antibacterial activity and can be used in gene therapy. The sequences
XX CC can also be used in the diagnosis and treatment of bacterial infections,
XX CC particularly S. epidermidis infections. The sequences can be used to
XX CC screen for compounds able to interfere with the S. epidermidis life
XX CC cycle or inhibit S. epidermidis infection.
XX CC N.B. The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from the
XX CC USPTO web site.
XX SQ Sequence 448 AA;

Query Match 20.9%; Score 543.5; DB 23; Length 448;
Best Local Similarity 30.5%; Pred. No. 1.7e-37;
Matches 137; Conservative 103; Mismatches 182; Indels 27; Gaps 9;

QY 85 IGLSVHTAPVEMREKLAELPRAIQELTSLNHIEEAAVLSTCNRMIEIYVVALSWNRG 144
DB 6 ISINHRTADVTLEQVAFRDALRLAHEDLYETKAILLENVILSTCNRTVEIVAVDQVHTG 65

QY 145 IREVDWMKSKGIPASELREHLFILRSSDATRHLFEVSGAGLSLVGEGQILAQVKQV 204
DB 66 RYIQRFLARSGFGEVDVDDIKDMSEVKVGDVAVEHLRVTSGLSIVLGETQILQOMRDAF 125

QY 205 RSGQSGGLGNIDRMFKDAITAGKRVSRSETNISGSAVSVSSAAVELALMKLPKSEALSA 264
DB 126 FLAQTGTGTGIFNHLFKQAIFAKKAHSETDIADNAVSVSYAAVELAKKVFGLK--SK 183

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 12:56:14 ; Search time 37 Seconds
(without alignments)
2669.117 Million cell updates/sec

Title: US-10-018-902-4

Perfect score: 2606

Sequence: 1 MATTTSATTAATAATAATKAP.....DMEKALIEQKIKAKVENTON 531

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09D_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	483.5	18.6	463	10	US-09-738-626-3970
2	428.5	16.4	449	12	US-09-882-227-508
3	427	16.4	569	15	US-10-156-761-12273
4	143	5.5	1258	15	US-10-156-761-10395
5	139.5	5.4	975	12	US-10-080-608A-19
6	139.5	5.4	975	12	US-10-370-685-108
7	138	5.3	1725	15	US-10-037-182-12
8	138	5.3	1786	10	US-09-938-275-7
9	138	5.3	1786	15	US-10-037-182-10
10	137.5	5.3	1979	15	US-10-205-823-419
11	133	5.1	2563	11	US-09-836-705-46
12	130	5.0	975	12	US-10-094-749-2035
13	129	5.0	648	12	US-10-032-585-7540
14	128	4.9	869	9	US-09-815-242-5230
15	128	4.9	862	9	US-09-815-242-12526

16 128 4.9 999 15 US-10-156-761-14584
17 124.5 4.8 1178 15 US-10-128-714-8240
18 124.5 4.8 1566 15 US-10-156-761-14473
19 124 4.8 856 9 US-09-815-242-11134
20 124 4.8 945 9 US-09-745-763-191
21 124 4.8 1000 12 US-10-012-697-1540
22 124 4.8 1711 10 US-09-771-161A-219
23 124 4.8 1711 10 US-09-771-161A-220
24 119 4.6 970 15 US-10-078-531-4
25 119 4.6 1765 15 US-10-037-182-8
26 119 4.6 1786 9 US-09-873-676-113
27 119 4.6 1786 10 US-09-938-275-6
28 119 4.6 1786 15 US-10-037-182-6
29 118 4.5 928 12 US-10-080-608A-23
30 118 4.5 928 12 US-10-370-685-112
31 118 4.5 1289 10 US-09-712-363-259
32 118 4.5 2285 10 US-09-932-183A-2
33 117.5 4.5 931 12 US-10-032-585-7600
34 117.5 4.5 1192 12 US-10-053-662A-32
35 117.5 4.5 1441 15 US-10-128-714-8123
36 117.5 4.5 4999 10 US-09-976-059-14
37 117 4.5 969 15 US-10-078-531-8
38 117 4.5 971 15 US-10-078-531-7
39 117 4.5 1008 15 US-10-078-531-2
40 116.5 4.5 1388 15 US-10-146-473-82
41 116 4.5 971 15 US-10-078-531-6
42 116 4.5 974 15 US-10-156-761-10499
43 116 4.5 1039 15 US-10-128-714-3123
44 116 4.5 1441 12 US-10-412-897-3
45 116 4.5 1641 14 US-10-017-216-5

ALIGNMENTS

RESULT 1

US-09-738-626-3970
; Sequence 3970, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OGIAL, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patentin ver. 3.0
; SEQ ID NO 3970
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-3970

Query Match 18.6%; Score 483.5; DB 10; Length 463;
Best Local Similarity 27.7%; Pred. No. 1.1e-31;
Matches 124; Conservative 102; Mismatches 186; Indels 35; Gaps 9;

QY 81 TIAVIGLSVHTAPVEMREKLAVALAEELWPRAIQELTSLNHIEEAALVSTCNRMIEIYVVALS 140

```
Db 2 SVLIWGMHRSAPVSELLERLSMDDSVRGTTQALLGRASLSALIVSTCNRELEVYVTSS 61
Qy 141 WNRGIREVDMWMSKSGIPASERLHFLIRSSDATRHLFEVSAGLDSVLGEGQILAQV 200
Db 62 PHTGVNDVVEVLHEASGVDIETLRGLVYRYADAAAEHMLVVTSGLDNVLGEGQIIQV 121
Qy 201 KOVVRSGQNGGLKNDIMFKDAITAGKRVSTETNISSGAVSVSSAAVELALMKL---P 257
Db 122 RTAYQAAEYSGVGPALHSLTQTALHTGKRVHSETAIDDAGASMYSFAVDRLVQMGDLS 181
Qy 258 KSEA--LSAR--MLLIGAGKMGKLVIKHLVAKGCKVWVNRVSVERVDAIREEMKDIEIVR 315
Db 182 EAEAPLSGKTALVLCAGAMSSLAATHLGRAGISNLIIMANRTLERAEHLAEHSLGAVPAE 241
Qy 316 PLSDMYQAAA--EADVTTSTASSTSLFAKEHAELPPVSDTMGCVRLFVDISVPRNVSA 373
Db 242 VVEYDQASAVNRVDLVVSATGDDF-----TVKPEDIPEGASLMVLVDSMPERDID 293
Qy 374 CVSEVGAARVNVVDLKEVVEANKEDRLRKAMEAQTIITEELRRFEAWRDSLETPTIKK 433
Db 284 ACADUPGVDLVNIERLHKASREGSGMAPSEEEALAVREELDSFTSQRIINDIVPAVSA 353
Qy 434 LRSYADRIASLEKCLQKV-----GEDALTKMRRRAIEELSTGIWNKLLHGLQLH 485
Db 354 LRRQAASVGSDELRLRQAPGISEVNGE--VEKTVR-----VVDKLLHEPTVRV 403
Qy 486 R--CDGSDSRTLDETLNHNHMRMFS 510
Db 404 KELAARSGSISYDSALQELFGLLESAS 430

RESULT 2
US-09-882-227-508
; Sequence 508, Application US/0988227
; Publication No. US20030158396A1
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Oomenen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding No. US20030158396A1el Helicobacter Polypeptides in the
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 06132/047002
; CURRENT FILING DATE: US/09/882,227
; PRIOR FILING DATE: 2001-06-15
; PRIOR FILING DATE: US 08/902,615
; PRIOR FILING DATE: 1997-07-29
; NUMBER OF SEQ ID NOS: 638
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 508
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-882-227-508

Query Match 16.4%; Score 428.5; DB 12; Length 449;
Best Local Similarity 28.8%; Pred. No. 3.7e-27;
Matches 127; Conservative 81; Mismatches 186; Indels 47; Gaps 9;

Qy 94 VEMREKLAVABELPRAIQEL--TSLNH--IEEAAVLSTCNRMIEYVW--ALSWNGI 145
Db 22 LEMREKLAIN--ATLKEFLQTIKNCPNKECWLSTCNRMIEYVW--ALSWNGI 78

Qy 146 REVVDWMSKSGIPASERLHFLIRSSDATRHLFEVSAGLDSVLGEGQILAQV 205
Db 79 NALLKILAQNKMVSVDLEKCVLMNTDSSAVHHVSVSCSSLDLSLVGETQITGMKAYK 138

Qy 206 SQNGSGGLKNDIMFKDAITAGKRVSTETNISSGAVSVSSAAVELALMKLPKSEALSAR 265
Db 139 FAFEEKCSKDLTLLHFAKCAAKVRNLTGISQGVSISSVAKEALNIFEKERIKDK 198
```

```
Qy 266 MLLIGAGKMGKLVIKHLVAKGCKVWVNRVSVERVDAIREEMKDIEIVRPLSDMYQAAA 325
Db 199 ALVIGLGEMALVIGKHLNKKQFEALILGRNAKPEDFIKELEPKKVSFQNTENLWAYIN 258
Qy 326 EADVVTSTAS-----ETSLFAKEHAELPPVSDTMGCVRLFVDISVPRNVSA 375
Db 259 EYELFCATSPHFTVQNRMLKETIFR-----RPFEDLAVPRNIEKPV 301
Qy 376 SEVGAARVNVVDLKEVVEANKEDRLRKAMEAQTIITEELRRFEAWRDSLETPTIKK 435
Db 302 --LDNIFLYSDLEPMVRENVENRQSRMRAYEIVGLATMFEFYQWIOSLEVEPIKDLR 359
Qy 436 SYADRIASLEKCLQKVGEDALTKMRRRAIEELSTGIWNKLLHGLHLCRDCGSDSRTL 495
Db 360 ELARISAKELQKALK---RYVPKEYENNETIKLHNAFNTFLHNPTTALKKNAQK---- 412
Qy 496 DETLENHMLNRMESLDMEKA 516
Db 413 ESDVLVGAIKNLFNLDKNSA 433

RESULT 3
US-10-156-761-12273
; Sequence 12273, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12273
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12273

Query Match 16.4%; Score 427; DB 15; Length 569;
Best Local Similarity 24.4%; Pred. No. 7.1e-27;
Matches 137; Conservative 94; Mismatches 184; Indels 146; Gaps 14;

Qy 81 TIAVIGLSVHTAPVEMREKLAVAEELPRAIQELTSLNHIEEAAVLSTCNRMIEYVVALS 140
Db 2 SLLVVGSHRSAPVSVLERATLTADAQVKLLQDTAAEPATEAAVLATCNRIELYADV 61

Qy 141 WNRGIREVDMWMSKSGIPASERLHFLIRSSDATRHLFEVSAGLDSVLGEGQILAQV 200
Db 62 FHAGVAELSTLLAQHSGVGLLEELTPYLVVHYEDRAVHHLFSVACGLDSMVVGEQILQI 121

Qy 201 KOVVRSGQNGGLKNDIMFKDAITAGKRVSTETNISSGAVSVSSAAVELALMKLPKSE 260
Db 122 KDALATAQELHTAGRLNLDLFQOALRTGKRAHSETGIDRAGQSL-----VTFGLQLSAGT 177

Qy 261 ALSA-----RMLLIGAGKMGKLVIKHLVAKGCKVWVNRVS--VERVDAIREEMKDIEI 312
Db 178 AVEAWKKGKALVIGAGSMSSLAATLARAGVAEVIANRTPDRAERLAQILTEGDDTDV 237

Qy 313 VYR--PLSDMYQAAAEEADVVTSTAS-----ETSLEA-- 342
Db 238 LARAVPMADVPAELTRADVAVSCTGATGLVLTAEAAVAEAGRTGTPVAVREETPASAG 297
```

QY 343 -----KEH----- 345
Db 298 GLAPAGTDEGCPLDLSAVOGATGFSVMGEAAVAGMDAATLEQRAAWDRGTVDRRDSRT 357
QY 346 -----AEALPPVSDTM-GVRL----- 380
Db 358 PEVEAELITALAATAVAGRLPERRPBPVVEAPRAPALALLDLAMPDRIDAAVHRLG 417
QY 381 ARVNVDDLKVEV-----ANKEDRLRKAMEAQTIITEELRRFEAWRDSLETVPTIKLRS 436
Db 418 VRLVDIESLAESADAPWAADVDLVRR-----IVADEVAAFGAQAARAAHITPTVVALRT 471
QY 437 YADRIRASELSEKLOKQGEDALTCKMRRRAIBELSTGIVNKLHGP-----LOHLRCDCGDSR 493
Db 472 MAADVANEIARLDGRL-----PGLDEKQRCGEITQTVRRVVDKLLHAPTAVRVKQLAAEPGAG 529
QY 494 TLDETLENNHAINRMFSLDME 514
Db 530 YAD-----ALRTLFDLDPE 543

RESULT 4

US-10-156-761-10395

; Sequence 10395, Application US/10156761

; Publication No. US20030119018A1

; GENERAL INFORMATION:

; APPLICANT: OMURA, SATOSHI

; APPLICANT: IKEDA, HARUO

; APPLICANT: ISHIKAWA, JUN

; APPLICANT: HORIKAWA, HIROSHI

; APPLICANT: SHIBA, TADAYOSHI

; APPLICANT: SAKAKI, YOSHIYUKI

; APPLICANT: HATTORI, MASAHIRA

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-262

; CURRENT APPLICATION NUMBER: US/10/156,761

; CURRENT FILING DATE: 2002-05-29

; PRIOR APPLICATION NUMBER: JP 2001-204089

; PRIOR FILING DATE: 2001-05-30

; PRIOR APPLICATION NUMBER: JP 2001-272697

; PRIOR FILING DATE: 2001-08-02

; NUMBER OF SEQ ID NOS: 45109

; SEQ ID NO 10395

; LENGTH: 1258

; TYPE: PRT

; ORGANISM: Streptomyces avermitilis

US-10-156-761-10395

Query Match

Best Local Similarity 5.5%; Score 143; DB 15; Length 1258;

Matches 124; Conservative 76; Mismatches 214; Indels 172; Gaps 24;

QY 45 DAAGVEAQAQAVAKAASVAA-----LEQPKISADRYMKERSTIAV-IGLSVHTAP 93
Db 43 ESADLGYAEQLLENAQIQADQLRADAEELSQAQTORILQEHAEQAARLQAEHTEA 102
QY 94 VENEKEL-----AVABELPRAIQELTSLNHIEE-----AAV 125
Db 103 VTRRQQLDQELAEERQTVESHVNVENVAQOLRARSEQQARRL--VEESRAEABQALAA 160
QY 126 LSTCNRM-----EIVVVALSWNRGIREVVDMWSKSGIPASELRHLFTLRS 172
Db 161 RABAEERIAAEARQRLNTDAESARAEAEALLRRARTDAERLLNASTQAQEAHTAEQLRS 220
QY 173 SDATRLHFEVSAGLSDLVIGEGQILAQVKQVRSQNSGGLGKNIIDRMFKDAITAGKVR 232
Db 221 STAT-----ESDA-----ARQASELSRTAEQRMTEADAALREAR 255
QY 233 SETNISSGAVSVSAAVELAMKLPKSEALSARMLLIGAKMGKLVTKHLVAKGCKVVV 292
Db 256 AE-----ADKVLTEAKEAAKALASAESANEQRT-----RTAKEQVARLVSEATKEAT 304
QY 293 VNRSEVRVDAIREMKDIEIVYRPLSDMYQAAAEADVFTSTASET-SLFAKEHAELPP 351

Db 305 TKAAEEVVA-----DAXAEAKILTAAADKARSLETAETASQLTK 345
QY 352 VSDTMGGVRLFVDISVPRNVSACVSEVGAARVYVDDLKBYVVEANKK-DRLR----- 402
Db 346 AARTAEVDNLKASBEAKSTTKAAATEE--AERIRR-----EASTRADRLRAEAHDIAE 395
QY 403 -----KAMBAQTI-ITEELRRFEAWRDSL--ETVPTIKLRSADRIRASELEKC 449
Db 396 QLKGTAKDDTKYRAKTVELQSEARRLUGDAEQLRADAAVEGERIRSEARREAVQOIIEBA 455
QY 450 LQKVGEHDALTK-----KMR-RAIEELSTGIVNKLHGPQLHLCDCGSD 491
Db 456 -AKTAEELLAKAKADADELRTAATTDEKVTETALERTATT--LRQABETLERTNEAR 512
QY 492 SRTLDETLENNHAIN-----RMFSLDMEXAIIIEQKIKAKVETQ 530
Db 513 HRA--EAVEQSEAIKEDAERAARELRDDSERAIERARAEADELTR 556

RESULT 5

US-10-080-608A-19

; Sequence 19, Application US/10080608A

; Publication No. US20030198956A1

; GENERAL INFORMATION:

; APPLICANT: Makowski, Lee

; APPLICANT: Hyman, Paul

; APPLICANT: Williams, Mark

; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES

; FILE REFERENCE: 8471-010-999

; CURRENT APPLICATION NUMBER: US/10/080,608A

; CURRENT FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 180

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 19

; LENGTH: 975

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

US-10-080-608A-19

Query Match

Best Local Similarity 5.4%; Score 139.5; DB 12; Length 975;

Matches 117; Conservative 104; Mismatches 230; Indels 129; Gaps 22;

QY 2 ATTTSATTAATAAAATTAKPRGSSAL-----CORVAGGRRRSVVVRCDAAGVEA 51
Db 405 AACTAAAEAAALAAQORTALANNSASVAVNEQARLATECERL-----YQQLDDKDEI 455
QY 52 QAQA-VAKAASVAALAEQFKISAD---RYMKERSTIAVIGLSVHTAPVEMREKLAVAEEL- 106
Db 456 NQDSQYAEQLKEQVMEQEELIANARREYETLQSEMARIQQENESAKSEVKEVLQALELA 515
QY 107 --WPRATQELTSLNHIEEAAVLSTCNRMETVWVALSWNRGIREVVDMWSKSGIPASELR 164
Db 516 VNYDQKSEIDNKNKIDALNBELOQKQSVFNAASTELQQLKDMSSHQKKR-----ITEML 571
QY 165 EHLFILRSSDATHLFEVSAGLSDLVIGEGQILAQVKQVRSQNSGGLGKN-----I 217
Db 572 TNL-----RDLEGVG--QAIAPGESSI--DLKMSALAGTDAKVEEDFTMARLFI 618
QY 218 DRMFKDAITAGKVRVS-ETNISSGAVSVSAAVELALMKLPKSEALSARMLLIGAKMGK 276
Db 619 SKMTEAKNTAQCNSMNETQQADSNKKIHEYKDLGEYRL----- 658
QY 277 LVIKHLVAKGCKKVVVNRSEVRVDAIREEMKDIEIVYRPLSDMYQAAAE--ADVFTST 334
Db 659 -----LISQHEARKMLQBSMRBAENKKTLEBOIDSLRECEAKLKAHEH 703
QY 335 ASSTSLFAKEHAELPPVSDT-MGVRLFVDISVPRNVSACVSEVGAARVYVDDLKKEV- 392
Db 704 VSAVNAEERQRAELSRMFPDSQMDLR-----EAHTRQVSELRDEI-AAKHQHEMDKMDVH 758
QY 393 -----VEANKEDRLRKAMEAQTIITEELRRFEAWRD-----SLETVP 429

QY 475 NKLHGLPQLHRCDSGSDRTLDTELNNHMLNRMFSLDMEKAIIEQKIKAKVEKTON 531
Db 1612 EKVVSVKQNA---DDVKTLTGELDEKY--KKVESLIAQKTEESADARRKAEILLQN 1663

RESULT 8
US-09-938-275-7
; Sequence 7, Application US/09938275
; Patent No. US20020111309A1
; GENERAL INFORMATION:
; APPLICANT: Gerardo Castillo
; APPLICANT: Alan Snow
; TITLE OF INVENTION: Therapeutic and Diagnostic Applications
; FILE REFERENCE: PROTEO.P03
; CURRENT APPLICATION NUMBER: US/09/938,275
; CURRENT FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: Mus Musculus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Swissprot P02469
; DATABASE ENTRY DATE: 1989-07-01
US-09-938-275-7

Query Match 5.3%; Score 138; DB 10; Length 1786;
Best Local Similarity 19.0%; Pred. No. 0.04;
Matches 125; Conservative 101; Mismatches 243; Indels 188; Gaps 27;

QY 28 CORVAG-GRRRS-----GVVRCDAAGVEAQA--QAVAKAASVAALAEQPKI 70
Db 1103 CQCMFGGRTCECOELFWGDPDVECRACDPRGIETPCQDQSTGQCVCVGEVGP RC 1162

QY 71 SADRYMKERSTIAVIGLSVHTAPVEMREKLAVAEELWPRAIOELTSLNH--IEEAVALST 128
Db 1163 -----DKCTGYSGVFPDCTPCHQ-----CFALMDAIIIGELNTRTHKFLKAKAKI 1209

QY 129 CNRMEIYVVALSWNRGIREVVDMMSKK-----SGIPASELRHFLPIL-----RSSD 174
Db 1210 SGVIGPY-----RETVDSEKKVNEIKDILAQSPAEPKNIIGILFEEAEKTKD 1259

QY 175 ATRHLFEVSAGLDSL-----LGEGQI-----LAQKVQVRSQNSGGIG 214
Db 1260 VTERKMAQVEVKLTDTASQNSTAGELGALQAEASLDKTVKELAEQLEFIKNSDIQ GAL- 1318

QY 215 KNIDRMFKDAITAGKRVRSSETNISGAVSVSAA---VELALMK-----LPKSEALSARM 267
Db 1319 DSITKYFQMSLEAEKRVNASITDPNSTVEQSALTDRVEDLMLERESPFKEQOEQARLL 1378

QY 268 LIGAGKMGKLV-----KHLVAKGCKKVVVNR 295
Db 1379 DELAGKQLSLDLASAAQMTCTPGADCSECCGPNCRDTBEKKCGCGGLVTVAH 1438

QY 296 SV--ERVDATRE--EMKDIIVYRPLSDMYQAAAEA-----DVVFTTASSETSL----- 340
Db 1439 SAWQKAMDFDRDLVSALAEVQLSKVSEAKVRADKQNAQDVLTKTNATKEKVDKSN 1498

QY 341 -----FAKEHAELPPVSDTMGGVRLFVDISVPRNVSVACVSEGAARVNVVDL 389
Db 1499 DLRLNIKQIRNFLTSDSADLSIEAVANVLKSGNASTPQQLNLTEDI--RERVETLSQV 1557

QY 390 KEVVEANKEDRLRKAMEAQTIIITELRRFPAWRDSLETPTIKKLSYADRIRASELEK 449
Db 1558 EVILQQAADIAR-----AELLLEAKRASKSATDKVTADVMVKEALEAEKQAQA-AEKA 1612

QY 450 LQKVGED-----ALT-----KQMRRAIEEL-----STGIV 474
Db 1613 IKQADEDIQGTQNLTLTSEETAASEETLTNASQRIKSLERNVEELKRAQAQNSGEAYI 1672

QY 475 NKLHGLPQLHRCDSGSDRTLDTELNNHMLNRMFSLDMEKAIIEQKIKAKVEKTON 531

Db 1673 EKVVSVKQNA---DDVKTLTGELDEKY--KKVESLIAQKTEESADARRKAEILLQN 1724

RESULT 9
US-10-037-182-10
; Sequence 10, Application US/10037182
; Publication No. US20030044899A1
; GENERAL INFORMATION:
; APPLICANT: Trygsvaasen, Karl
; APPLICANT: Doi, Masayuki
; APPLICANT: Thyboll, Jill
; TITLE OF INVENTION: Recombinant Laminin 10
; FILE REFERENCE: 99-274-F
; CURRENT APPLICATION NUMBER: US/10/037,182
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,449
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/279,282
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-037-182-10

Query Match 5.3%; Score 138; DB 15; Length 1786;
Best Local Similarity 19.0%; Pred. No. 0.04;
Matches 125; Conservative 101; Mismatches 243; Indels 188; Gaps 27;

QY 28 CORVAG-GRRRS-----GVVRCDAAGVEAQA--QAVAKAASVAALAEQPKI 70
Db 1103 CQCMFGGRTCECOELFWGDPDVECRACDPRGIETPCQDQSTGQCVCVGEVGP RC 1162

QY 71 SADRYMKERSTIAVIGLSVHTAPVEMREKLAVAEELWPRAIOELTSLNH--IEEAVALST 128
Db 1163 -----DKCTGYSGVFPDCTPCHQ-----CFALMDAIIIGELNTRTHKFLKAKAKI 1209

QY 129 CNRMEIYVVALSWNRGIREVVDMMSKK-----SGIPASELRHFLPIL-----RSSD 174
Db 1210 SGVIGPY-----RETVDSEKKVNEIKDILAQSPAEPKNIIGILFEEAEKTKD 1259

QY 175 ATRHLFEVSAGLDSL-----LGEGQI-----LAQKVQVRSQNSGGIG 214
Db 1260 VTERKMAQVEVKLTDTASQNSTAGELGALQAEASLDKTVKELAEQLEFIKNSDIQ GAL- 1318

QY 215 KNIDRMFKDAITAGKRVRSSETNISGAVSVSAA---VELALMK-----LPKSEALSARM 267
Db 1319 DSITKYFQMSLEAEKRVNASITDPNSTVEQSALTDRVEDLMLERESPFKEQOEQARLL 1378

QY 268 LIGAGKMGKLV-----KHLVAKGCKKVVVNR 295
Db 1379 DELAGKQLSLDLASAAQMTCTPGADCSECCGPNCRDTBEKKCGCGGLVTVAH 1438

QY 296 SV--ERVDATRE--EMKDIIVYRPLSDMYQAAAEA-----DVVFTTASSETSL----- 340
Db 1439 SAWQKAMDFDRDLVSALAEVQLSKVSEAKVRADKQNAQDVLTKTNATKEKVDKSN 1498

QY 341 -----FAKEHAELPPVSDTMGGVRLFVDISVPRNVSVACVSEGAARVNVVDL 389
Db 1499 DLRLNIKQIRNFLTSDSADLSIEAVANVLKSGNASTPQQLNLTEDI--RERVETLSQV 1557

QY 390 KEVVEANKEDRLRKAMEAQTIIITELRRFPAWRDSLETPTIKKLSYADRIRASELEK 449
Db 1558 EVILQQAADIAR-----AELLLEAKRASKSATDKVTADVMVKEALEAEKQAQA-AEKA 1612

QY 450 LQKVGED-----ALT-----KQMRRAIEEL-----STGIV 474
Db 1613 IKQADEDIQGTQNLTLTSEETAASEETLTNASQRIKSLERNVEELKRAQAQNSGEAYI 1672

QY 475 NKLHGLPQLHRCDSGSDRTLDTELNNHMLNRMFSLDMEKAIIEQKIKAKVEKTON 531

```
Db 1673 EKVVSVKQNA---DDVKKTLGDGELDEKY--KKVESLIAQKTESADARRKALLQN 1724

RESULT 10
US-10-205-823-419
; Sequence 419, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Womsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; SOFTWARE: FastSeq for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 455
; SEQ ID NO 419
; LENGTH: 1979
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-419

Query Match
Best Local Similarity 5.3%; Score 137.5; DB 15; Length 1979;
Matches 102; Conservative 87; Mismatches 213; Indels 81; Gaps 18;

Qy 87 LSVHTAPVEMREKLAVALAEELMPRAIQEL-TSLNHIEAAVLST-CNRMEIYVVALSWNRG 144
Db 600 VSIQENLELKEHIRONEEELSIRNELMQSLNQDSNSNFKDTLLKEREAEVRNLKQNL 659
Qy 145 IREVVDMWKSXGIPASERLHLFLRSDATRHLEFVSAGLDSVLVGEQILAQVKQV 204
Db 660 ELEQLNENLKKVAFDVKNENKL-VLACEDVRHOLEECLAGNQLSLEKNTIVETLK--M 716
Qy 205 RSGNSGGGLGKNIIDRMFKDAITAGRKVRSETNISSGAVSVSAAVELA--LMKLPKSEA 261
Db 717 EKGEIEALCWAKKLEELAEANKYEKTI BELN-----ARNLNTSAQLHEHLKLNKKO 772
Qy 262 LSARMLLTGAGKMG-----KLVI-----KHLVAKGCKKVVVNRSVRVDIAIREMK 308
Db 773 MEIAELKKNIEQMDTHKEDVLSLSSLEQKQLTQLINKKEIFIEKLKERSKQLQELD 832
Qy 309 DIEIVYRPLSDMYQAAADVVFTSTASETSLF-----AKEHAELPPVSDTMGGVRLF 362
Db 833 KYSOALRKNELIRQTIIEKDRSLGSMKEENHLOEELERLEBOQSRTAPVAD----- 884
Qy 363 VDISVPRNVSAVSEVGAARVYVVDLKEVVEAN-----KEDRLRKAMEAQTIITTEL 415
Db 885 -----PKTLDLDS-VTEL-ASEVSQLNTIKEHEEBEIKHKHQLIEIQNQSCKMLQSLQSQK 937
Qy 416 RRFEAWRDSLETVPITKILRSYADRIIRASELEKCLQKVGEDALTKQKRRRAIEELSTGIVN 475
```

```
Db 938 KEMDEFRIQHE-----QMNATHTQLFLEK---DBEIKSLQKTIQIKT----- 977
Qy 476 KLHGLPLQHLRCGSD-----SRTLDETLEN---MHALNRMFSLDMEKAIIEQIKAKVEK 528
Db 978 -QLHEERQDIQTDSNDFIQETKVQSLNIENGSEKHDLSKAETERLVKGIKERELEIKLAN 1036
Qy 529 TQN 531
Db 1037 EKN 1039

RESULT 11
US-09-836-705-46
; Sequence 46, Application US/09836705
; Publication No. US20030078395A1
; GENERAL INFORMATION:
; APPLICANT: Abe, Yuki
; APPLICANT: Ono, Chiho
; APPLICANT: Yoshikawa, Hiroji
; TITLE OF INVENTION: Genes from a Gene Cluster
; FILE REFERENCE: 01149/HG
; CURRENT APPLICATION NUMBER: US/09/836,705
; CURRENT FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: JP 2000-116591
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: JP 2000-117458
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 46
; LENGTH: 2563
; TYPE: PRT
; ORGANISM: Penicillium citrinum
US-09-836-705-46

Query Match
Best Local Similarity 5.1%; Score 133; DB 11; Length 2563;
Matches 80; Conservative 59; Mismatches 163; Indels 78; Gaps 13;

Qy 8 TTAATAAATTAKPRGSSSALCORVAGGRRRSRGVVRCDAAAGVEAQAQAAVAAALEQ 67
Db 1976 TTAVIALVTTAKLRGGERVLHSGA-GGVGQAAIILSQLAGAEVFTAGTQAKRDFVGDK 2034
Qy 68 FKISADRYMKERSTIAVIGLSVHTAPVEMREKL-AVABELMPRAIQELTSLNHIEAAV- 125
Db 2035 FGINPDHIFSSRNDLFDVGIRKAYTGGVHVVLNSLAGQLLQASDFCDMAEFGRFVEIGKK 2094
Qy 126 -LSTCNMEIY-----VVALSMNRGIREVVDMWKSXGIPASERLHLFLRSD 174
Db 2095 DLEQNSRLDMLPFRTRDVSPTSIDLLSWQAKSEEV-----SE 2131
Qy 175 ATRHLFEVSAGLDSVLVQ-----EQILAQVKVVRSCQNSGGGLCKNIDRMFKDAITAGK 229
Db 2132 ALNHVTKL---LETKAIGLIGPIQHSLSLNIEKAPRTMQSQHVGVVVVNSGDEL----- 2184
Qy 230 RVRSETNISSGAVSVSAAVELALMKLPKSEALSARMMLLIAGKMGKLVIKHLVAKGCKK 289
Db 2185 -----VPVGDGGFSLKLK--PDSSYLVAQ-----GLGGIGKQKICQWLVDHGAKH 2226
Qy 290 VVVNRSVERVDATREEMKDIIEIVYRPLSDMYQAAAEADVFTSTASETSLFAKEHAEL 349
Db 2227 LIILSRSAKAPFTITSQNCQCAVILHACDI-----SDQDV-----TKVLRICEEHAH 2275
Qy 350 PPVSDTMGGVRLFVDISVPR 369
Db 2276 PPIRGIIQGAWLKALLSR 2295

RESULT 12
US-10-094-749-2035
; Sequence 2035, Application US/10094749
; Publication No. US20030219741A1
```



```
RESULT 14
US-09-815-242-5230
; Sequence 5230, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5230
; LENGTH: 869
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5230

Query Match          4.9%; Score 128; DB 9; Length 869;
Best Local Similarity 21.6%; Pred. No. 0.094;
Matches 119; Conservative 85; Mismatches 197; Indels 150; Gaps 28;

Qy 47 AGVEAQ-----AQAVAKAASVAALAEQFKISADRYNKE-----RSTIAVIGLSVH 90
Db 72 ASVEGDNIQYQYISQAN-----QLITKAESYKMEYEDEYISMEHILRSAMDIDQTTKH 126
Qy 91 TAPVEMREKLAVAEELWPRAIOELTSLNHIEEAVALSTCNRMELIYVVALSWNRGIREVVD 150
Db 127 -----YINKKVEIKDI-----IKKVRGGNHV-----TSQNPVNYEALA--KYGRDLVE 169
Qy 151 WMSKKSQIPA-----SELREHLFILRSSDATRHLFEVSAGLDSLVLGEGQILAQVKQVVR 206
Db 170 EVRQGMDFVIGRDEEIRNTIRILSRKTKNPVLIGEPGVKTAIVEG--LAQ--RIVKK 225
Qy 207 QNSGGLKNDIRMFKDAITAGKVRSETNISSGAVSVSSAAVELALMKLPKSEA----- 261
Db 226 DVPESLDKTVFELDLSALVAGAKYRGFEERLKAV-----LKEVKESDGRILF 275
Qy 262 LSARMLLIGAGKM-GKL-----VIKHLVAKG-----CKKVVVNRVERV 300
Db 276 IDEIHLVAGKTDGADGNWMLPMLARGELHICIGATTLINEYREYIEKOSALERRFOKV 335
Qy 301 DAIREMKDIEIVRPISDMYQ-----AAAAADVFTSTASSETSLFAKEHA- 346
Db 336 AVSEPDVEDTISILRGLKERYEVHGVRIQDRALVAAAE-----LSDRYITDRFLPKAI 390
Qy 347 -----EALPPVSDTMGGVRLFDVDSVPRNV-----SACVSEVCAARVYNVDLKEVEA 395
Db 391 DLVDQACATRTNGMSGNPTELD-QVNRVRVMOLETESALKNESDNASKORLQELQEEAN 449
Qy 396 NKE-----DLRLKAMEAQTIITEELRRFEAWRDSLETVTPTIKLRSVADRIRAS--ELE 447

Db 450 EKEKQAAALQSRVSESEKIANVLQKRAQLDESRLAQDAQTNNLNKAAELQYGTIPQLE 509
Qy 448 KCLQKV-----GEDALTCKMERRAI---EELS-----TGIV-----NKLH-G 480
Db 510 KELRELEDNFDQEQEGD--TDRMIREVVUTDEIGDIVSQWTGIPVSKLVETEREKLLHLS 567
Qy 481 PLOHLRCDGSD 491
Db 568 DILHKRVVGD 578

RESULT 15
US-09-815-242-12526
; Sequence 12526, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12526
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12526

Query Match          4.9%; Score 128; DB 9; Length 882;
Best Local Similarity 21.6%; Pred. No. 0.096;
Matches 119; Conservative 85; Mismatches 197; Indels 150; Gaps 28;

Qy 47 AGVEAQ-----AQAVAKAASVAALAEQFKISADRYNKE-----RSTIAVIGLSVH 90
Db 85 ASVEGDNIQYQYISQAN-----QLITKAESYKMEYEDEYISMEHILRSAMDIDQTTKH 139
Qy 91 TAPVEMREKLAVAEELWPRAIOELTSLNHIEEAVALSTCNRMELIYVVALSWNRGIREVVD 150
Db 140 -----YINKKVEIKDI-----IKKVRGGNHV-----TSQNPVNYEALA--KYGRDLVE 182
Qy 151 WMSKKSQIPA-----SELREHLFILRSSDATRHLFEVSAGLDSLVLGEGQILAQVKQVVR 206
Db 183 EVRQGMDFVIGRDEEIRNTIRILSRKTKNPVLIGEPGVKTAIVEG--LAQ--RIVKK 238
Qy 207 QNSGGLKNDIRMFKDAITAGKVRSETNISSGAVSVSSAAVELALMKLPKSEA----- 261
Db 239 DVPESLDKTVFELDLSALVAGAKYRGFEERLKAV-----LKEVKESDGRILF 288
Qy 262 LSARMLLIGAGKM-GKL-----VIKHLVAKG-----CKKVVVNRVERV 300
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 12:53:43 ; Search time 21 Seconds

(without alignments)
1669.860 Million cell updates/sec

Title: US-10-018-902-4

Perfect score: 2606

Sequence: 1 MATTSATTAAAAATAKPP.....DMEKAIIEQIKAKVETQN 531

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

- Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	557.5	21.4	435	4	US-09-252-991A-30562
2	543.5	20.9	448	4	US-09-134-001C-3055
3	498	19.1	429	4	US-09-328-352-4643
4	206.5	7.9	341	4	US-09-198-452A-759
5	139.5	5.4	975	4	US-09-914-259-19
6	132.5	5.1	871	4	US-09-134-001C-3979
7	121	4.6	1939	4	US-09-310-187A-1
8	119.5	4.6	1886	4	US-08-938-105-3
9	119	4.6	1196	1	US-08-144-121-4
10	119	4.6	1196	2	US-08-735-893-4
11	118	4.5	928	4	US-09-914-259-23
12	118	4.5	2285	4	US-09-308-375-2
13	117	4.5	573	4	US-09-252-991A-18744
14	116.5	4.5	1388	4	US-09-572-191-2
15	116.5	4.5	1388	4	US-09-723-262-2
16	116.5	4.5	1388	4	US-09-723-219-2
17	115.5	4.4	3248	1	US-08-353-700-1
18	115.5	4.4	3248	5	PCT-US95-16216-1
19	115	4.4	2101	1	US-08-466-390-4
20	115	4.4	2101	1	US-08-470-950-4
21	115	4.4	2101	1	US-08-467-781-4
22	115	4.4	2101	1	US-08-195-487-4
23	115	4.4	2101	2	US-08-483-924-4
24	115	4.4	2101	3	US-09-452-294-1
25	115	4.4	2101	5	PCT-US93-06160-4
26	114.5	4.4	439	4	US-09-252-991A-31097
27	114.5	4.4	941	4	US-09-336-447A-9

28	113.5	4.4	2482	1	US-08-328-254-6	Sequence 6, Appli
29	113	4.3	756	3	US-09-085-199B-9	Sequence 9, Appli
30	112.5	4.3	614	4	US-09-252-991A-31412	Sequence 31412, A
31	112	4.3	748	4	US-09-252-991A-31491	Sequence 31491, A
32	112	4.3	1713	3	US-08-600-982-24	Sequence 24, Appl
33	112	4.3	1713	5	PCT-US94-10261A-24	Sequence 24, Appl
34	111.5	4.3	592	2	US-08-736-770-6	Sequence 6, Appli
35	111.5	4.3	592	4	US-09-702-705-1809	Sequence 1809, Ap
36	111.5	4.3	592	4	US-09-736-457-1809	Sequence 1809, Ap
37	111.5	4.3	5215	3	US-09-105-537-2	Sequence 2, Appli
38	111	4.3	388	4	US-09-252-991A-21907	Sequence 21907, A
39	111	4.3	10182	4	US-09-134-001C-3159	Sequence 3159, Ap
40	110.5	4.2	696	4	US-09-252-991A-16965	Sequence 16965, A
41	110	4.2	332	4	US-09-252-991A-29420	Sequence 29420, A
42	110	4.2	477	4	US-09-328-352-5428	Sequence 5428, Ap
43	110	4.2	2188	4	US-09-328-352-7763	Sequence 7763, Ap
44	109.5	4.2	467	4	US-09-252-991A-21258	Sequence 21258, A
45	109.5	4.2	1024	4	US-09-562-737-41	Sequence 41, Appl

ALIGNMENTS

RESULT 1

US-09-252-991A-30562
; Sequence 30562, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30562
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30562

Query Match	21.4%	Score	557.5	DB	4	Length	435
Best Local Similarity	32.3%	Pred. No.	2.3e-43				
Matches	142	Conservative	91	Mismatches	187	Indels	19
Gaps	7						
Qy	76	MKERSTTAVIGLSVH--TAPVEMREKLAVAEELWPRAIOELTSLNHIEEAAVLSTCNRME	133				
Db	8	LAORSCNFAITAGINHKHTASVAVRERVAFTPEQMVEALQQLCRLLTSREAAILSTCNRS	67				
Qy	134	IYVVALLSWNRGIREVDWMSKSGIPASELREHLFILRSSDATRHLFEVSGAGLSVLGE	193				
Db	68	LYLEI--DHTADDVLAWLADYHRLTDELRACAVHQDSDVAVHMRVAVSGLSGMLVGE	125				
Qy	194	GQILAQVQVVRSGQSGGLGKNDIRMFKDITAGTKEVRSETNITSSGAVSVSSAAVELAL	253				
Db	126	PQILGQKMSAYAVAREAGTVGLLGRLFQATFTSTAKTVTDTAIGENPVSVFAVSLA-	184				
Qy	254	MKLPKSEALSARMLLTGAGKMGKLVIKHLVAKGKKVVRVNRVSRVVDATREEMKDIEIV	313				
Db	185	-KQIFSDLRHRSQALLIGAGETITLVARHLFEOGVKRVIVANRTLERASLLAEQFGAHL	243				
Qy	314	YRPLSDMYQAAARADVVTSTASETSLFAKEHAELPPVSDTWGGVTLFVDISVPRNVA	373				
Db	244	---LSPIPELANSDIVISTASQLPILGKGAVER--ALKQKHKPMFMYDIAVPRDIEP	298				
Qy	374	CVSEVGAARYNVDDLKVEVVEANKEDRLRKAMEAQTIITELRRFEAWRDSLETPTIKK	433				
Db	299	EVGELDVIYISVDDLHEVVAENLKSQGNAAQAEELVGSVAFEFQRLRELAADVLR	358				


```

Db      492  TEGNLEKAAELQYGTIGTIOLEKELQEFEEAFQDETGED--SERMIREVVVSDEIGDIVSQW 549
Qy      471  TGI-VNKLKH-----GPLQHLRCDGSD 491
      |||:|:|:
Db      550  TGIPVSKLVETEREKILSLSDILHKRVVGQD 580

RESULT 7
US-09-310-187A-1
; Sequence 1, Application US/09310187A
; Patent No. 6358751
; GENERAL INFORMATION:
; APPLICANT: Benichou, Gilles
; APPLICANT: Fedoseyeva, Eugenia
; TITLE OF INVENTION: Involvement of Autoantigens in Cardiac
; TITLE OF INVENTION: Graft Rejection
; FILE REFERENCE: UCSP-090
; CURRENT APPLICATION NUMBER: US/09/310,187A
; CURRENT FILING DATE: 1999-05-12
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1939
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-310-187A-1

Query Match      4.6%; Score 121; DB 4; Length 1939;
Best Local Similarity 17.3%; Pred. No. 0.098;
Matches 106; Conservative 118; Mismatches 248; Indels 142; Gaps 22;

Qy      34  GRRRRSGVRCDAAGVAAQAQAVAKAA--SVAALAEQFKISADRYMKERSTTAVICLSVHT 91
Db      1038 GSLEQEKVYMDLERAKRKLEGLDKLTQESIMDNLNDKIQLEEKKKKE-----FDINQ 1091
Qy      92  APVEMREKLAVAEELWIPRAIQELTSLNHIEEAAVLSTCNRMIEIYVVALSMNRGIREVVDW 151
Db      1092 QNSKIEDEQALALQOKLKENQARIIEELEEAERTAKRVEKRLSDLSRELEISER 1151
Qy      152 MSXKSG-----IPASELREHLFILRSSD--ATHLFEVSAGL-----DSLVLGE----- 193
Db      1152 LEEAGGATSVQIEMNKKREAFQKVRDLDEEATLQHEATAALRKKHADSVAELGEQIDN 1211
Qy      194  -----GQILAQVQKVVRSGNSGGLGKNI--DRMFKDAITAGKVRVS 233
Db      1212 LQRVQKLEKESEFFKLELDDVTNSMEQIKAKANLEKVSITLEDQANEYVKLEEAQRS 1271
Qy      234  ETNISSGAVSVSSAAVELALMKLPKSEALSARMLLIGAKMGKL-----VIKHLVAK 285
Db      1272 LNDFTTQRAKLQTEGELA-RQLEKEALISQL-----TRGLSYTQQMEDLKRQLEEE 1324
Qy      286  GCKKVVVNRSVVERVDAIREEMKDIEIVYRPLSDMYQAAAEADVVFSTASSETSFP----- 341
Db      1325 GKAKNALAH-----ALQSAHGHCDLRLREQYEEETAKAELQVLKANSSEVAQWRTKY 1377
Qy      342  -----AKEHAELPPVSDTMGGVR-----LFPDISVP 368
Db      1378 ETDAIORTEEEBAKKLQALQLODAEAVEAVNAKCSLEKTKHRLQNEITDLNVVERS 1437
Qy      369  RNVSACVSEVGAARVYVNDL-----KEVNEANKEDRLRKAMEAQTIITEELRRFEAWRDS 424
Db      1438 NAAAAALDK---KQRFNFKILAEWKQKYBESQSELESSEKARSLSSTELFKLNAYEES 1493
Qy      425  LETVPTIKK-----LRSYADRI-----RASELEKCLKVGSDALTQKWRRAIELST 471
Db      1494 LEHLFTFKRENKNLQBEISDLTEQLSGGKNVHELEKVRQKQLEVEKL---ELQSLAESEAA 1551
Qy      472  GIVN----KLHGLPQHLRCDGSGSRSTLDETLENMHALNR-----MFLSDMEKAIIE 519
Db      1552 SLEHEGKILRAQLAEFNOIKAEIERKLAEKDEMEQAKRNHORVVDLSLQTSLDATTSRN 1611
Qy      520  Q--KIKAKVEKTON 531

```

Db 1377 --LMVDVRSNAAAALDK---KQNFQKILAENKQYEESSQSESSQKARSLSLSTEL 1430
 QY 415 LRRFARWDSLETPTIKK-----LRSYADRI-----RASELEKCLQKVGEDALTKK 461
 Db 1431 FKLKWAYESLSHLTFKRNKVLQBEISDLTEQJGEGGKNVHELEKIRKQLEVEK--E 1488
 QY 462 MRAIEELSTGIVN---KLLHGPLQHLRCGSDSRTSLDTLENMHALNR-----MF 509
 Db 1489 LQSALEAEASLEHEEGKILRAQLEFNQIKAEIERKLAEKDEMEQAKRNHLRVVDSLQT 1548
 QY 510 SLDMEKALIEQ--KIKAKVEKTON 531
 Db 1549 SLDAETRSRNEALRVKKXMEGDLN 1572

RESULT 9

US-08-144-121-4
 ; Sequence 4, Application US/08144121
 ; Patent No. 5610031
 ; GENERAL INFORMATION:
 ; APPLICANT: Burgeson, Robert E.
 ; APPLICANT: Wagman, David W.
 ; TITLE OF INVENTION: BLK CHAIN OF LAMININ AND METHODS OF USE
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; CITY: BOSTON
 ; STATE: Massachusetts
 ; COUNTRY: United States
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/144.121
 ; FILING DATE: 27-OCT-1993
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Myers, Paul L. 35,965
 ; REGISTRATION NUMBER:
 ; REFERENCE/DOCKET NUMBER: (MGP-0780.0) MGP-021
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 227-7400
 ; TELEFAX: (617) 227-5941
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1196 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; FRAGMENT TYPE: internal
 ; FEATURE:
 ; NAME/KEY: Domain
 ; LOCATION: 1..250
 ; FEATURE:
 ; NAME/KEY: Domain
 ; LOCATION: 251..437
 ; FEATURE:
 ; NAME/KEY: Domain
 ; LOCATION: 438..807
 ; FEATURE:
 ; NAME/KEY: Domain
 ; LOCATION: 808..840
 ; FEATURE:
 ; NAME/KEY: Domain
 ; LOCATION: 841..1196

Query Match 4.6%; Score 119; DB 1; Length 1196;
 Best Local Similarity 19.4%; Pred. No. 0.069;

Matches 115; Conservative 87; Mismatches 209; Indels 182; Gaps 24;
 QY 28 CORVAG--GRRRS-----CVRCDAAAGVTAQA--QAVAKAASVAALAEQFKI 70
 Db 513 CQCPGFGGRTCEBQELFWGDGPDVCEKACDPRGIETPQDQSTGQCVCVGEVGPGR 572
 QY 71 SADRYMKERSITIAVIGLSVHTAPVEMREKLAVALPRAIOELTSLNH--TEEAVALST 128
 Db 573 -----DKCTRGVSGVPPDCTPCHQ-----CPALMDVIIABLTNRTHRFLEKAKALKI 619
 QY 129 CNRMBIYVALSWNRGIREVVDMMSKK-----SGIPASELRHLFIL----- 170
 Db 620 SGVIGPY-----RETVDSEVERKVKSEIKDILAQSPAAEPLKKNIGLNPPEAEKLIK 669
 QY 171 -----RSSDATHLPFVSAGLSLVLEGQI-----LAQVKQVTSQNSGGILG 214
 Db 670 VTEMMAQVEVKLSDTTSQSNSTAKELDSLQTEAESLQNTVKELAEQLEFIKNSDTRGAL- 728
 QY 215 KNIDRMFKDAITAGKRVRSNETNISGAVSVSSA---AVELALM-----KLPKSEALSARML 267
 Db 729 DSITYFQMSLEAEERVNASTTEPNSTVEQSALMRDRVEDVMMERESQFKKQESQARLL 788
 QY 268 LIGAGKMGKLYI-----KHLVAKGCKKVVVNR 295
 Db 789 DELAGKLQSLQLSAAAEWTCGTPGCASCSETECGPNCRCTDEGERKCGGPGGLVTVAH 848
 QY 296 SVERVDAIREEMKDIIEIVRPLSDMYQAAADVVFTTASSETSLFAKEHAELPPVSDT 355
 Db 849 NA-----WQKAMDLD-----QDVLASALAEVQL-SKQVSEAKLRADEAKQSAE----- 890
 QY 356 MGVRLFDVISPRNVSAVCSEVGAARVYVDDLKEVVEANKEDRLKAMEAQTITTEL 415
 Db 891 -----DILKTNAT-----KEMDKSNEELRNLIKOIRNFLTOD- 924
 QY 416 RRFEAWRDSLETPTIKLRSYADRIIRASELEKCLQKVGEDALTKMERRAIEELSTGIVN 475
 Db 925 --SADLDSIEAV-----ANEVLKMEPSTPQQL--QNLIEDIERVESLSQ----- 966
 QY 476 KLLHGPLQHLRCGSDSRTSLDTLE--NMHALRMFSLDMKALIEQKIKAKV 526
 D.. 967 --VEVILQHSAAADIARAEMLEEARAKSKSATDVKVTADMYKVALEAEAKQV 1017

RESULT 10
 ; Sequence 4, Application US/08735893
 ; Patent No. 5914317
 ; GENERAL INFORMATION:
 ; APPLICANT: Burgeson, Robert E.
 ; APPLICANT: Wagman, David W.
 ; TITLE OF INVENTION: BLK CHAIN OF LAMININ AND METHODS OF USE
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; CITY: BOSTON
 ; STATE: Massachusetts
 ; COUNTRY: United States
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/735,893
 ; FILING DATE: 18-OCT-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/144,121
 ; FILING DATE: 27-OCT-1993
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:

NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: (MCH-0780.1) MGP-021DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1196 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: Domain
LOCATION: 1..250
FEATURE:
NAME/KEY: Domain
LOCATION: 251..437
FEATURE:
NAME/KEY: Domain
LOCATION: 438..807
FEATURE:
NAME/KEY: Domain
LOCATION: 808..840
FEATURE:
NAME/KEY: Domain
LOCATION: 841..1196
US-08-735-893-4

Query Match 4.6%; Score 119; DB 2; Length 1196;
Best Local Similarity 19.4%; Pred. No. 0.069;
Matches 115; Conservative 87; Mismatches 209; Indels 182; Gaps 24;
QY 28 CORVAG-GGRRS-----GVVRCDAAGVEAQA--QAVAKAASVAALQPKI 70
DB 513 CQCMFGFGRTCEQQLFWGPDVECRACDPRGIETPCQDQSTGQCVCVGGPRC 572
QY 71 SADRYKERTSTAVTGLSVHTAPVEMREKLAVALAEELPRAIQELTSLNH--IEEAVALST 128
DB 573 -----DKCTRGYGVFPDCTPCHQ-----CFALMDVITAEIENRTHRFLEKAKALI 619
QY 129 CNRMEIYVVALSWNRGIREVVDWMSK-----SGIPASLREHLFIL-----170
DB 620 SGVIGPY-----RETVDSEKVKSEIKDILAQSPAEPPLKNIGLFEAEKLIKD 669
QY 171 -----RSSDTRHLFEVSAGLDSLVLGEGOI-----LAQVKQVVRSGQNGSLG 214
DB 670 VTEMAQVEVKLSLDTTOSNSTAKELDSLQTEAESLDNTVKELARQLEFIKNSDIRGAL- 728
QY 215 KNIDRMFKDAITAGKRVRSSETNISSGAVSVSSA--AVEALAM-----KLPKSEALSARM 267
DB 729 DSITYKQNSLEAEERVAATTEPNSTVEQSALMRDRVEDVMWRESOFKEKQEQARLL 788
QY 268 LIGAKMGKGLVI-----KHLVAKGCKKVVVYNR 295
DB 789 DELAGKLSLDLSAAAEWTCGTPPGCASCECTCGGPNCRKTDEGERKCGPGCGGLVYAH 848
QY 296 SVERVDAREEMKDIIVYRPLSDMYQAAAEADVFTSTASTSIFAKEHAELPPVSDT 355
DB 849 NA-----WQKAMDLD-----QDVLASALAEVEQL-SKWVSEAKURADEAKQSAE- 890
QY 356 MGVRLFDVDSVPRNVACSVEGAARVVYVNDLKEVVEANKEDRLRKAMEAQTIITEEL 415
DB 891 -----DILLKTNAT-----KEKMDKNEEURNLIKQIRNFLTQD- 924
QY 416 RRFEMRWDSLETVPITIKKLSYADIRASELEKCLKQKVGEDALTKMRRRAIEELSTGIW 475
DB 925 ---SADLSIEAV-----ANEVLKMEPSTPQQL--QNLTEDIRERVESLSQ----- 966
QY 476 KLLHCPLOHLRCGSDSRTLDTLE--NMHALNRMFSLDMEKAIIEQIKAKV 526
DB 967 ---VEVILQSAADTARAEMLLSEAKRASKSATDVKVTDAMVYKEALEEAEKAV 1017

RESULT 11

US-09-914-259-23
; Sequence 23, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914.259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 928
; TYPE: PRT
; ORGANISM: Neurospora crassa
US-09-914-259-23

Query Match 4.5%; Score 118; DB 4; Length 928;
Best Local Similarity 20.3%; Pred. No. 0.057;
Matches 113; Conservative 95; Mismatches 200; Indels 150; Gaps 24;
QY 7 ATTAAAAATTAKPRGSSSALCQRVAGGRRRRSGVVRCDAAAGVEAQAQAAKAAASVAAL 66
DB 391 AITPSKASATTARPSTPSRLPES-----RAETPAISDRAGTCTSLP 431
QY 67 QPKISADRYMK-----ERSTIAVIGLSVHTAPVEMREKLAVALAEELPRAIQELTS 116
DB 432 LDKDEREFLERENELODQIAEKESIAA-----AAERQLRE---TKSEL--IALKDHS 480
QY 117 LNHTEAAVLSTCNRMEIYVVALSW-NRGIREVVDWMSKSGIPASELREHLFILRSS-- 173
DB 481 KLGENEELISESENEFQMLERLAFENKEAQITIDGLKDN-----SELTAEDELVEKQML 536
QY 174 DATRHLFEVSAAGLSVLGEGOIILAQVKQVVRSGQNGSLGNIDRMFKDAITAGKRVRS 233
DB 537 DMKSAKETSAVLDE--KEKKAKAEKMAKMGAGDLSGDVPSDNERAVADAIA---QLDA 590
QY 234 ETNLSGAVSVSSAAVELALMKLPSE-----ALSARMLLIGAKMGKGLVIKHLVA 284
DB 591 LFEISSAGDAIPPEIDIKALREKLVETQGFVROAELSFSAASSDAEARKRAELEAR-LEA 649
QY 285 KGCKKVVVVRNSVERVDAIREEMKDIEIVYRPLSDMYQAAAE-----ADVFTSTAST 338
DB 650 LQOEHEELLSRNLTEAD--KEEVK--ALLAKSLSDKSQVVELVEQLKADIALKN--SET 703
QY 339 SLFAKEHAELPPVSDTWGGVRLFDVDSVPRNVACSVEGAARVVYVNDLKEVVEANKE 398
DB 704 -----EHLKAL-----VDDLQRRVKAGGA 722
QY 399 DRLRKAMEAQTIITEELRRFEAMRDSLETVPITIKKLSYADIRASELEKCLKQKVGEDAL 458
DB 723 G---VAMANGTVOQQLAEPDMVKSL-----MRDLQNRCEV--VELEISLDETREQ- 770
QY 459 TKMRRRAIEELSTGIYNKLLHGPLOHLRCGSDSRTLDTLENMHALNRMF-----SLDM 513
DB 771 -----YNNVLS--SNNAQQKKAFLEARNLEQLTQVORQLVEQNSALKK 813
QY 514 EKAIIEQIKAKVKTQN 531
DB 814 EVAIERKLMARNERIQS 831

RESULT 12

US-09-308-375-2
; Sequence 2, Application US/09308375
; Patent No. 6300117
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.

; TITLE OF INVENTION: Proteases From Gram-Positive Organisms
; FILE REFERENCE: GC394-PCT
; CURRENT APPLICATION NUMBER: US/09/308,375
; CURRENT FILING DATE: 1999-05-14
; EARLIER APPLICATION NUMBER: EP9719636.4
; EARLIER FILING DATE: 1997-09-15
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 2285
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-09-308-375-2

Query Match 4.5%; Score 118; DB 4; Length 2285;
Best Local Similarity 20.7%; Pred. No. 0.24;
Matches 107; Conservative 86; Mismatches 189; Indels 136; Gaps 21;

QY 62 VAALBQFISADRYMKRSTIAVIGLSVHTAPVEMREKLAVAEELWPRAIOELTSLNHIE 121
DB 914 IADDDYKVAADK-----AKQSMKLAQSDIESGNKV 945
QY 122 EAVLSTCNREIYVVALSWNGIREVVDWMSKSGIPASELREHLFILRSSDATRHLFE 181
DB 946 KDSVLSIANAYSSIDISNTLTSISDVVYKMLKDDLDPEELEKF-----SSSLGKLOEK 1000
QY 182 VSAGLDSVLGEGQILAQVQVVRSGQSGGLGKNIIDRMFKDAITAGKRVRSVETNISGA 241
DB 1001 MOKALDS---GDEKAFDNKKDLQSLLETYSKSDSSIDVFKMSFD-----KAQRNKKDGD 1052
QY 242 VSVSSAAVELALMKLPKSEALSARMLLIGAGKMGKLVIKHLVAKGCKKVVVNVRSVERVD 301
DB 1053 KSLSSV-----KSEVGDGLGETLAEGAGNEAEDFGKLL-----KEALQAN-SVDDIK 1096
QY 302 AIREEMKD-----IEIVYR-----PLSDMYQAAAEADVFTSTASVETSLPAK 343
DB 1097 AAIKEMSDAMQFDSQVYDNLGDIENNTKDQVAPLNDLLEKMAEGKSI---SANEANTLIQ 1153
QY 344 EHAELPPVSTMGVRLFDVDSVPRNVSAVSEVGAARVYVVDLKEVVE-ANK-----397
DB 1154 KOKELAQAISIEGVVKNRD-----EVIKORKVKLDAYNDMVYVSNKMKTE 1201
QY 398 EDRLRKAMEAQTIITEELRRF-----EAWRDSLETPTIKKLSYADRIASEL-- 446
DB 1202 VVNAIKTLNADTLRIDSLKKLRKERKLDMSAEISDLSE---VKSINNVD--AKKELK 1255
QY 447 --EKCLKVG-----EDALTKOMRAIEELSTGIVNKLHGLPQHLRCDGS-- 490
DB 1256 LEEKHLOPGGYSNSQIEAMQSVKSALESYISASEEATSTQEMNK-----QALVEAGTSL 1309
QY 491 DSRTLDLETENHMLNRMFSLDMKAEIIEQKIKAKVEK 528
DB 1310 ENWTDQBEKANEETKTSYVVDKYKEALE-KVNAEIDK 1346

RESULT 13

US-09-252-991A-18744
; Sequence 18744, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18744
; LENGTH: 573

; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18744

Query Match 4.5%; Score 117; DB 4; Length 573;
Best Local Similarity 19.1%; Pred. No. 0.033;
Matches 101; Conservative 75; Mismatches 182; Indels 170; Gaps 20;

QY 6 SATTAATAAATYAKPRGSSSALCORVAGGRRRRSVVRCDAAGVBAQAOVAKAASVAAL 65
DB 157 SATSARVLSLAASDP-----SAAQQLSYGSDSRQFCGMR-----EVINQLDEMEEAAAADG 208
QY 66 EQFKISADRYMKRSTIAVIGLSVHTAPVEMREKLAVAEELWPRAIOELTSLNHEEARV 125
DB 209 EASSALGERHRWQVALVAFGL-----230
QY 126 LSTCNREIYVVALSWNGIREVVDWMSKSGIPASELREHLFILRSSDATRHLFEVSAG 185
DB 231 -----LVCLS-----LVLFPGVLTFRPLQRLQR 254
QY 186 LDSVLGEGQILAQVQVVRSGQSGGLGKN-----IDRMFKDAITAGKRVRSVETNISGA 241
DB 255 LEEIANGDGL--RVRLEVTSRDEPGRGLSAPNAPFLDKL-QPLIREVGRVGTGEVADSAGS 311
QY 242 VSVSSAAVELALMKLPKSEALSARMLLIGAGKMGKLVIKHLVAKGCKKVVVNVRSVER-- 299
DB 312 LAGMTAAND-----KLINSEHASVDQVSTAATOMSSAV--HEVARNAQSAQAQVADARRQA 365
QY 300 -----VDAREEMKQIEIVYRPLSDMYQAAAEADVFT--STASVETSLFAKEHA 346
DB 366 REGANVVEATIEVIRLOAQEVESSESIEQLAQETASIDAVLTWIKGTAEQTNLLALNAA 425
QY 347 EALPPVSTMGVRLFDVDSVPRNVSAVSEVGAARVYVVDLKEV-----VEANKEDR 400
DB 426 IEAARAGEQGRGFVAVD-----EVRALAARTQDSTKDIQARIERLQAGVQNA 473
QY 401 LRKAMEAQTIITEELRRFPEARWDSLETPTIKKLSYADRIASELEKCLKVKGEDALTK 460
DB 474 VR-AMQSGSLKA-----RDSVE-----RAAGVGVGLAATG-DAVGR 507
QY 461 KVRRAIBELSTGIVNKLHGLPQHLRCDGSRTLDLETENHMLNRM 508
DB 508 -----INDLAAQIAS-----ACE-EQSRVIDEIRNISIVREL 539

RESULT 14

US-09-572-191-2
; Sequence 2, Application US/09572191
; Patent No. 6355466
; GENERAL INFORMATION:

; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6355466el motor proteins and methods for
; FILE REFERENCE: 1017
; CURRENT APPLICATION NUMBER: US/09/572,191
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1388
; TYPE: PRT
; ORGANISM: Human
US-09-572-191-2

Query Match 4.5%; Score 116.5; DB 4; Length 1388;
Best Local Similarity 19.0%; Pred. No. 0.15;
Matches 115; Conservative 92; Mismatches 199; Indels 199; Gaps 26;

QY 93 PVEMREKL--AVAEELWPRAIOELTSL-----NHIEAAVLSVCNRMEL 134
DB 704 PPEWNEQAFPAISEEL--RTVQEQMSALQAKLDEEHNKLNKQOQVDDKLEHHSTQMGELF 761

Result No.	Score	Query			DB	ID	Description
		Match	Length				
1	2379	91.3	536	1	HEM1_ORYSA	O48674 oryza sativa	
2	2313	88.8	527	1	HEM1_HORVU	Q42843 hordeum vulg	
3	2225	85.4	535	1	HEM3_HORVU	O85796 hordeum vulg	
4	2113	81.1	465	1	HEM2_HORVU	Q96563 hordeum vulg	
5	1784.5	68.5	552	1	HEM1_CUSCA	P93111 cucumis sativ	
6	1770.5	67.9	543	1	HEM1_ARATH	P42804 arabidopsis	
7	1680	64.5	530	1	HEM2_ARATH	P42804 arabidopsis	
8	1573	60.4	542	1	HEM2_CUSCA	P49295 cucumis sativ	
9	909.5	34.9	428	1	HEM1_ANASP	P49295 cucumis sativ	
10	904	34.7	427	1	HEM1_SNY33	P28463 anabaena sp	
11	893.5	34.3	426	1	HEM1_SNYEL	P28463 anabaena sp	
12	872.5	33.5	432	1	HEM1_CVAPA	P48077 cyanophora	
13	646	24.8	419	1	HEM1_VIBVU	Q9df88 vibrio vuln	
14	643.5	24.7	443	1	HEM1_HELMO	Q9zgg6 heliobacill	
15	637.5	24.5	415	1	HEM1_NEIMB	P56934 neisseria m	
16	632.5	24.3	415	1	HEM1_NEIGO	P56934 neisseria m	
17	628.5	24.1	415	1	HEM1_NEIMA	P56934 neisseria m	
18	628	24.1	418	1	HEM1_VIBPA	Q9zhd6 neisseria g	
19	620	23.8	419	1	HEM1_VIBCH	Q9zhd6 neisseria g	
20	616.5	23.7	459	1	HEM1_BACHO	Q97rn5 vibrio para	
21	613.5	23.5	437	1	HEM1_PAEWA	Q9kg24 vibrio chol	
22	597	22.9	416	1	HEM1_SHEON	Q9k8f8 bacillus ba	
23	596.5	22.9	426	1	HEM1_XANAC	O69108 paenibacill	
24	594.5	22.8	426	1	HEM1_XANCH	O69108 paenibacill	
25	590.5	22.7	432	1	HEM1_XANCP	Q9pn04 xanthomonas	
26	588	22.6	457	1	HEM1_OCEIH	P42808 xanthomonas	
27	577.5	22.2	420	1	HEM1_YERPE	Q9pc67 xanthomonas	
28	573.5	22.0	418	1	HEM1_SCOLI	Q9ezx9 versinia pe	
29	573	22.0	455	1	HEM1_BACSU	P13580 escherichia	
30	570.5	21.9	418	1	HEM1_SCOL5	P16618 bacillus eu	
31	570	21.9	435	1	HEM1_DESVM	Q9f103 escherichia	
32	569.5	21.9	418	1	HEM1_SALTY	Q9k452 desulfot vibr	
33	561	21.5	426	1	HEM1_CHLTE	P13581 salmonella	
34	561	21.5	418	1	HEM1_CHLFE	O938t4 chlorobium	

Db 2 MASTTSATAGCAFAAAKTRRVVGGRRPCARVAAGRRRRSGVVVRCD-AGVEAQAQA 60
 QY 56 VAKAASVAALQFKISADRYMKERSTIAVIGLSVHTAPVEMREKLAVALAEELWPAIQBEL 115
 Db 61 VAKAASVAALQFKISADRYMKERSIAVIGLSVHTAPVEMREKLAVALAEELWPAISELT 120
 QY 116 SUNHIEEAVALSTCNRMETIYVVALSWNRGIREVVDWMSKKGIPASELREHLFILRSSDA 175
 Db 121 SUNHIEEAVALSTCNRMETIYVVALSWNRGLREVVDWMSKKGIPASELREHLFLMRDSDA 180
 QY 176 TRHLEFVSAGLSVLGEGQILAQVQVRSQNSGGGLGKNDIRMFKDIAITAGKVRSET 235
 Db 181 TRHLEFVSAGLSVLGEGEILAQVQVRSQNSGGGLGKNDIRMFKDIAITAGKVRSET 240
 QY 236 NISSGAVSVSSAAVELALMKLPKSEALSARMILLIGAGKMGKULVIKHLVAKGCKKVVVNR 295
 Db 241 NISSGAVSVSSAAVELALMKLPKSEALSARMILLIGAGKMGKULVIKHLVAKGCKKVVVNR 300
 QY 296 SVERVDATREEMKDIETVVRPLSDMYOAAEADVVFTSTASSETSLPAKEHEALPPVSDT 355
 Db 301 SVERVDATREEMKDIETVVRPLSDMYOAAEADVVFTSTASSETPLTKEHAELPAISDA 360
 QY 356 MGVRLFVDIISVPRNVSCVSEVGARVYVVDLKEVVEANKEDRLRKAMEAQTIITBEL 415
 Db 361 MGVRLFVDIISVPRNVSCVSEVGARVYVVDLKEVVEANKEDRLRKAMEAQTIITBEL 420
 QY 416 RRFEARWRSLEVTPIKILRSVADRIRASELEKCLQKVGEDALTAKMRRAIBELSTGIVN 475
 Db 421 RRFEARWRSLEVTPIKILRSVADRIRASELEKCLQKVGEDALTAKMRRAIBELSTGIVN 480
 QY 476 KLLHGPLOHLRCGSDSRDLTLENMHALNMFSLDMEKATIEOKIKAKVEKTON 531
 Db 481 KLLHGPLOHLRCGSDSRDLTLENMHALNMFSLDMEKATIEOKIKAKVEKSON 536

RESULT 2

HMAL_HORVU STANDARD; PRT; 527 AA.
 AC Q42843; Q42844;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Glutamyl-tRNA reductase 1, chloroplast precursor (BC 1.2.1.-) (GluTR).
 GN HEMA1.
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
 OC Triticeae; Hordeum.
 OX NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Klages;
 RX MEDLINE=96267576; PubMed=8696365;
 RA Bougri O., Grimm B.;
 RT "Members of a low-copy number gene family encoding glutamyl-tRNA
 RL reductase are differentially expressed in barley.";
 RL Plant J. 9:867-878(1996).
 RN [2]
 RP SEQUENCE OF 44-65, AND CHARACTERIZATION.
 RX MEDLINE=95045501; PubMed=7957167;
 RA Pontoppidan B., Kannangara C.G.;
 RT "Purification and partial characterisation of barley glutamyl-
 RL tRNA(Glu) reductase, the enzyme that directs glutamate to chlorophyll
 bioynthesis";
 RL Eur. J. Biochem. 225:529-537(1994).
 RN [3]
 RP 3D-STRUCTURE MODELING OF 76-460.
 RX MEDLINE=20058652; PubMed=10591107;
 RA Brody S.S., Gough S.P., Kannangara C.G.;
 RT "Predicted structure and fold recognition for the glutamyl tRNA
 RL reductase family of proteins.";
 RL Proteins 37:485-493(1999).
 CC -1- CATALYTIC ACTIVITY: Glutamyl-tRNA(Glu) + NADPH = glutamate-1-

CC semialdehyde + NADP(+) + tRNA(Glu).
 CC -1- PATHWAY: Porphyrin biosynthesis by the C5 pathway; first step.
 CC Involved in chlorophyll biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- SIMILARITY: Belongs to the glutamyl-tRNA reductase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; X92403; CAA63140.1; -;
 CC SBL; X86101; CAA60054.1; ALT_INIT.
 CC PDB; 1B29; 09-DEC-98.
 CC PDB; 1B61; 18-FEB-00.
 CC InterPro; IPR000343; Glutr.
 CC Pfam; PF00745; Glutr_dimer; 1.
 CC Pfam; PF05201; Glutr_N; 1.
 CC Pfam; PF05200; Glutr_NAD_bind; 1.
 CC TIGRPFAMs; TIGR01035; hemA; 1.
 CC PROSITE; PS00747; Glutr; 1.
 CC Porphyryn biosynthesis; Chlorophyll biosynthesis; Oxidoreductase;
 CC NADP; Transit peptide; Chloroplast; Multigene family; 3D-structure.
 CC TRANSIT 1 43
 CC CHAIN 44 527
 CC DOMAIN 286 289
 CC STRAND 79 91
 CC STRAND 96 99
 CC HELIX 104 125
 CC TURN 126 126
 CC STRAND 133 136
 CC TURN 139 140
 CC TURN 150 152
 CC TURN 155 159
 CC TURN 160 160
 CC TURN 170 182
 CC TURN 192 192
 CC TURN 193 197
 CC TURN 205 206
 CC TURN 209 213
 CC TURN 225 230
 CC TURN 233 251
 CC TURN 252 253
 CC TURN 259 260
 CC TURN 262 265
 CC TURN 269 281
 CC TURN 282 282
 CC TURN 285 288
 CC TURN 293 302
 CC TURN 303 303
 CC TURN 305 306
 CC TURN 309 309
 CC TURN 313 315
 CC TURN 316 318
 CC TURN 319 321
 CC TURN 325 327
 CC TURN 339 340
 CC TURN 345 345
 CC TURN 347 350
 CC TURN 351 352
 CC TURN 357 358
 CC TURN 365 365
 CC TURN 368 376
 CC TURN 377 378
 CC TURN 380 381
 CC TURN 384 385
 CC TURN 386 389
 CC TURN 390 390
 CC TURN 398 404
 CC TURN 405 405

FT HELIX 413 421
 FT TURN 422 422
 FT HELIX 423 436
 FT TURN 437 438
 FT HELIX 443 457
 FT TURN 458 459
 SQ SEQUENCE 527 AA; 57652 MW; D329E56450BE1165 CRC64;

Query Match
 Best Local Similarity 88.8%; Score 2313; DB 1; Length 527;
 Matches 468; Conservative 27; Mismatches 32; Indels 4; Gaps 2;

QY 1 MATTTATTAAAAATTAKPRSSSALCORVAGGRRRSGVVRCDAAAGVAAQAAVAKAA 60
 DB 1 MAGATSA-TAAAGAFAAAKARGPAAACPLWVAAGRRRSGVVRCDAGG---DAQAASKAA 56

QY 61 SVAALAEQKISADRYMKERSTIAVIGLSVHTAPVEMREKLAVAEELWPRATQELTSLNHI 120
 DB 57 SITALEQPKISADRYMKERKSSIAVIGLSVHTAPVEMREKLAVAEELWPRATSELTSLNHI 116

QY 121 EEAVALSTCNRMIEIYVVALSWNRGIREVVDMWSKSGIPASELREHLFLRSSDATRHLF 180
 DB 117 EEAVALSTCNRMIEIYVVALSWNRGIREVVDMWSKSGIPASELREHLFLRSSDATRHLF 176

QY 181 EVSAGLSLVLGEGQILAQVKQVVRSGNSGGLGNIDRMFKDAITAGKRVRSSETNISG 240
 DB 177 EVSAGLSLVLGEGQILAQVKQVVRSGNSGGLGNIDRMFKDAITAGKRVRSSETNISG 236

QY 241 AVSVSSAAVELALMKPKSEALSARMLLIGAGMGKLVKHLVAGCKKVVVVRNRSERV 300
 DB 237 AVSVSSAAVELALMKPKSEALSARMLLIGAGMGKLVKHLVAGCKKVVVVRNRSERV 296

QY 301 DAIREEMKDIEIVRPLSDMYQAAAEADVFTSTASSETSLFAKEHAEALPPVSDTMGGVR 360
 DB 297 DAIREEMKDIEIVRPLSDMYQAAAEADVFTSTASSETSLFAKEHAEALPPVSDTMGGVR 356

QY 361 LFVDISVPRNVSACVSEVGAARVYVNDLKEVVEANKEDRLRKAMEAQTIITEELRRPEA 420
 DB 357 LFVDISVPRNVSACVSEVGAARVYVNDLKEVVEANKEDRLRKAMEAQTIITEELRRPEA 416

QY 421 WRDSLETPTTKLRSYADRIASELEKLCQKVGEDALTKKRRRAIELSTGIVNKLHG 480
 DB 417 WRDSLETPTTKLRSYADRIASELEKLCQKVGEDALTKKRRRAIELSTGIVNKLHG 476

QY 481 PLOHLRCDSGDSRTLDLENNHNLNRMFSLDMKAEIIEQIKAKVEKTON 531
 DB 477 PLOHLRCDSGDSRTLDLENNHNLNRMFSLDMKAEIIEQIKAKVEKTON 527

RESULT 3

HMA3_HORVU STANDARD; PRT; 535 AA.
 AC O65796;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glutamyl-tRNA reductase 3, chloroplast precursor (EC 1.2.1.-) (Glutr).
 GN HMA3.
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Hordeum.
 OX NCBI_TaxID=4513;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=cv. Bonus; TISSUE=Root;
 RA Tanaka R., Yoshida K., Nakayashiki T., Tsuji H., Inokuchi H.,
 RA Okada K., Tanaka A.;
 RT "The third member of the hema gene family encoding glutamyl-tRNA
 RT reductase is primarily expressed in roots in Hordeum vulgare."
 RL Photosyn. Res. 53:161-171(1997).
 CC -1- CATALYTIC ACTIVITY: Glutamyl-tRNA(Glu) + NADPH = glutamate-1-
 semialdehyde + NADP(+) + tRNA(Glu).

CC -1- PATHWAY: Porphyrin biosynthesis by the C5 pathway; first step.
 CC Involved in chlorophyll biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- TISSUE SPECIFICITY: PRIMARILY EXPRESSED IN ROOTS.
 CC -1- SIMILARITY: Belongs to the glutamyl-tRNA reductase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D88383; BAA25168.1; -;
 DR HSR; T04402; T04402.
 DR PIR; Q42843; I829.
 DR InterPro; IPR000343; Glutr.
 DR InterPro; IPR000205; NAD binding.
 DR InterPro; IPR000594; ThiF domain.
 DR Pfam; PF00745; Glutr_dimer; 1.
 DR Pfam; PF05201; Glutr_N; 1.
 DR Pfam; PF05200; Glutr_NAD_bind; 1.
 DR TIGRFAMs; TIGR01035; hema; 1.
 DR PROSITE; PS00747; Glutr; 1.
 KW Porphyrin biosynthesis; Chlorophyll biosynthesis; Oxidoreductase;
 KW NADP; Transit peptide; Chloroplast; Multigene family.
 FT TRANSIT 1 CHLOROPLAST (POTENTIAL).
 FT CHAIN ? 535 GLUTAMYL-TRNA REDUCTASE 3.
 SQ SEQUENCE 535 AA; 58419 MW; B2830889AE6A3224 CRC64;

Query Match 85.4%; Score 2225; DB 1; Length 535;
 Best Local Similarity 84.0%; Pred. No. 1.4e-114;
 Matches 452; Conservative 38; Mismatches 38; Indels 10; Gaps 5;

QY 1 MATTTATTAAAAATTA---KPRSSSALCORVAGGRRR-SGVVRCDAAAGVAAQAA 55
 DB 1 MASTSTASATAGAFAAAGVYKPRG--SAACPRVPAGRGRLSCVVRCD-AGPGVPAQM 57

QY 56 VAKAASVAALAEQKISADRYMKERSTIAVIGLSVHTAPVEMREKLAVAEELWPRATQELT 115
 DB 58 AMAASVAALAEQKISADRYMKERKSSIAVIGLSIHTAPVEMREKLAVAEELWPRAVAE 117

QY 116 SLNHIEAVALSTCNRMIEIYVVALSWNRGIREVVDMWSKSGIPASELREHLFLRSSDA 175
 DB 118 NLNHIEAVALSPCNRMIEIYVVALSWNRGIREIYVDMWSKSGIPAVELREHLFMFRSDA 177

QY 176 TRHLFEVSAGLSLVLGEGQILAQVKQVVRSGNSGGLGNIDRMFKDAITAGKRVRSSET 235
 DB 178 TRHLFEVSAGLSLVLGEGQILAQVKQVVRSGNSGGLGNIDRMFKDAITAGKRVRSSET 237

QY 236 NISSGAVSVSSAAVELALMKPKSEALSARMLLIGAGMGKLVKHLVAGCKKVVVYNR 295
 DB 238 NISCGAVSVSSAAVELALMKPKSEALSARMLLIGAGMGKLVKHLVAGCKKVVYNR 297

QY 296 SVSRVDAIREMKDIEIVRPLSDMYQAAAEADVFTSTASSETSLFAKEHAEALPPVSDT 355
 DB 298 SVSRVDAIREMKDIEIVRPLSDMYQAAAEADVFTSTASSETSLFAKEHAEALPPVSDT 357

QY 356 MGGVRLPVDISVPRNVSACVSEVGAARVYVNDLKEVVEANKEDRLRKAMEAQTIITEEL 415
 DB 358 LGGVRLPVDISVPRNVSACVSDVGHARVYVNDLKEVVEANKEDRLRKAMEAQTIITEEL 417

QY 416 RRFEMARSDLETPTTKLRSYADRIASELEKLCQKVGEDALTKKRRRAIELSTGIVN 475
 DB 418 KRFEAMRSDMETPTTKLRSYADRVRASELKCLQKIGEDALTKKRRRSIEQLSTGIVN 477

QY 476 KLLHGPLOHLRCDSGDSRTLDLENNHNLNRMFSLDMKAEIIEQIKAKVE--KTON 531
 DB 478 RLLHGPLOHLRCDSGDSRTLDLENNHNLNRMFGLDTEKAVMEQIKTKVKEKQTON 535

RESULT 4

```

HMA2_HORVU
ID HMA2_HORVU STANDARD; PRT; 465 AA.
AC Q96563;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glutamyl-tRNA reductase 2 (EC 1.2.1.-) (Glutr) (Fragment).
GN HMA2.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Bonus;
RX MEDLINE=96267576; PubMed=8696365;
RA Bougri O., Grimm B.;
RT "Members of a low-copy number gene family encoding glutamyl-tRNA
RL reductase are differentially expressed in barley."
RL Plant J. 9:867-878(1996).
CC -!- CATALYTIC ACTIVITY: Glutamyl-tRNA (Glu) + NADPH = glutamate-1-
CC semialdehyde + NADP(+) + tRNA(Glu).
CC -!- PATHWAY: Porphyrin biosynthesis by the C5 pathway; first step.
CC Involved in chlorophyll biosynthesis.
CC -!- SUBCELLULAR LOCATION: Chloroplast.
CC -!- SIMILARITY: Belongs to the glutamyl-tRNA reductase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X86102; CAA60055.1; .
DR PIR; T05734; T05734.
DR HSSP; Q42843; 1B29.
DR InterPro; IPR000343; Glutr.
DR Pfam; PF00745; Glutr dimer; 1.
DR Pfam; PF05201; Glutr_N; 1.
DR Pfam; PF05200; Glutr_NAD bind; 1.
DR TIGRfams; TIGR01035; hema; 1.
DR PROSITE; PS00747; Glutr; 1.
KW Porphyrin biosynthesis; Chlorophyll biosynthesis; Oxidoreductase;
KW NADP; Chloroplast; Multigene family.
FT NON_TER 1
FT DOMAIN 224 227 POLY-VAL.
SQ SEQUENCE 465 AA; 51674 MW; E42A50A3E749CE6D CRC64;

Query Match
Best Local Similarity 81.1%; Score 2113; DB 1; Length 465;
Matches 418; Conservative 29; Mismatches 18; Indels 0; Gaps 0;

QY 67 QFKISADRYMKERSTIAVIGLSVHTAPVEMREKLAVALAEELWPRATQELTSLNHIEEAAVL 126
DB 1 QFKISADRYMEKSSIAVIGLSVHTAPVDMREKLAVALAEELWPRASELTSLNHIEEAAVL 60
QY 127 STCNRMETVYVVALSWNRGIREVDMWMSKSGIPASELREHLFILRSSDATTRHLFEVSAGL 186
DB 61 STCNRMETVYVVALSWNRGIREVDMWMSKSGIPASELREHLFMLRDSGATRHLLFEVSAGL 120
QY 187 DSLVLGEGQILAQVKQVVRSGNSGGLGKNIDRMFKDAITAGKRVSRSTNLSGAVSVSS 246
DB 121 DSLVLGEGQILAQVKQVVRSGNSGGLGKNIDRMFKDAITAGKRVSRSTNLSGAVSVSS 180
QY 247 AAVALAMKLPKSELSARMMLLIGAGKMGKLVKHLVAKGCKVWVNRVSRVDAITREE 306
DB 181 AAVALAMKLPKSELSARMMLLIGAGKMGKLVKHLVAKGCKVWVNRVSRVDAITREE 240
QY 307 MKQTEIVTRPLSDMYQAAAEADVFTSTASTSIFAKEHAALPPVSDTMGGVRLFVDIS 365
DB 1 MKQTEIVTRPLSDMYQAAAEADVFTSTASTSIFAKEHAALPPVSDTMGGVRLFVDIS 365

```

```

Db 241 MKQTEIVTRPLSDMYQAAAEADVFTSTASTSIFAKEHAALPPVSDTMGGVRLFVDIS 300
QY 367 VPRNVSAVSEVGARVYVNDLKSVEANKEDRLRKAMEAQTIITEELRPFVARDLSLE 426
Db 301 VPRNVGACLSQVEHARVYVNDLKSVEANKEDRVKAMEAQAIITQELKRFVARDLSLE 360
QY 427 TVPTIKLRSYADRIASELEKCLQKVGEDALTCKMRAIBELSTGIVNKLHGLOHLR 486
Db 361 TVPTIKLRSYADRIASELEKCLQKVGEDLNKKTREIBELSTGIVNKLHGLOHLR 420
QY 487 CDGSDSRITLDTLENMHALNMFSLDMKAIIEOKIKAKVKTON 531
Db 421 CDGSDSRITLDTLENMHALNMFSLDMKAIIEOKIKAKVKTON 465

RESULT 5
HMA1_CUGSA
ID HMA1_CUGSA STANDARD; PRT; 552 AA.
AC P93111;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glutamyl-tRNA reductase 1, chloroplast precursor (EC 1.2.1.-) (Glutr).
GN HMA1.
OS Cucumis sativus (Cucumber).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosida I; Cucurbitales; Cucurbitaceae; Cucumis.
OX NCBI_TaxID=3659;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Aonagajibai; TISSUE=Cotyledon;
RX MEDLINE=97088717; PubMed=8934625;
RA Tanaka R., Yoshida K., Nakayashiki T., Masuda T., Tsuji H.,
RA Inokuchi H., Tanaka A.;
RT "Differential expression of two hema mRNAs encoding glutamyl-tRNA
RL reductase proteins in greening cucumber seedlings."
RL Plant Physiol. 110:1223-1230(1996).
CC -!- CATALYTIC ACTIVITY: Glutamyl-tRNA (Glu) + NADPH = glutamate-1-
CC semialdehyde + NADP(+) + tRNA(Glu).
CC -!- PATHWAY: Porphyrin biosynthesis by the C5 pathway; first step.
CC Involved in chlorophyll biosynthesis.
CC -!- SUBCELLULAR LOCATION: Chloroplast (By similarity).
CC -!- TISSUE SPECIFICITY: PRIMARILY IN COTYLEDONS AND HYPOCOTYLS OF
CC GREENING CUCUMBER SEEDLINGS.
CC -!- SIMILARITY: Belongs to the glutamyl-tRNA reductase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D50407; BAA08910.1; .
DR PIR; T10186; T10186.
DR HSSP; Q42843; 1B29.
DR InterPro; IPR000343; Glutr.
DR Pfam; PF00745; Glutr dimer; 1.
DR Pfam; PF05201; Glutr_N; 1.
DR Pfam; PF05200; Glutr_NAD bind; 1.
DR TIGRfams; TIGR01035; hema; 1.
DR PROSITE; PS00747; Glutr; 1.
KW Porphyrin biosynthesis; Chlorophyll biosynthesis; Oxidoreductase;
KW NADP; Transit peptide; Chloroplast; Multigene family.
FT TRANSIT 1 2 CHLOROPLAST (POTENTIAL).
FT CHAIN ? 552 GLUTAMYL-tRNA REDUCTASE 1.
SQ SEQUENCE 552 AA; 60895 MW; 9C7A72AF24FF00F5 CRC64;

Query Match
Best Local Similarity 68.5%; Score 1784.5; DB 1; Length 552;
Matches 367; Conservative 74; Mismatches 86; Indels 27; Gaps 8;

```

QY 1 MATTTSATTA-----AAAATTAKPRGSSA-----LCORVAGGRR-----BSGV----- 41
 DB 1 MAVSTSFSAKUEALLFKASNSSTNLSHLPFCFKSIR--TRRILFQRTGVSSFTP 58
 QY 42 VRCDAAGVEAQAQ-----AVAKAASVAALBQFKISA-DRYMKERSTIAVLGLSVHTAPVEM 96
 DB 59 FRCLEASDVLVQNDIEDPPKSNLSALEBQKTSADVRYTKERSIIVIGLSIHTTPVEM 118
 QY 97 REKLAVALWELPRATOELTSLNHIIEAAVLSTCNMEIIVVALSNWRGIRREVVDWMSKS 156
 DB 119 REKLAIPAENWPRALGELCGLNHIIEAAVLSTCNMEIIVVALSQHRGVKTEVWMSKTS 178
 QY 157 GIPASELRHFLRSDATRHFLFEVSAGLSILVLGEGQILAQVKQVVRSGNSGGLGN 216
 DB 179 GIPVSEICQHRELLYNDATQHI FEVSAGLSILVLGEGQILAQVKQVVRSGVAGFGRN 238
 QY 217 IDRMFKDAITAGKVRSESTNIGSAGVSSAAVELALMKLPKSEALSARMLLIGAGKMK 276
 DB 239 ISGLFKHAIIVGKRVRTETNIAAGAVSVSSAAVELALMKLPKSEALSARMLLIGAGKMK 298
 QY 277 LVKHLVAKGCKVNVNRSVERVDAREEMKDIEIVRPLSDMYQAAABADVFTSTAS 336
 DB 299 LVKHLVAKGCKVNVNRSVERVDAREEMKDIEIVRPLSDMYQAAABADVFTSTAS 358
 QY 337 ETSLFAKHAELPPVSDTMGVGRVLFVDISVPRNVSACVSEVGAARVNVVDLKEVWAN 396
 DB 359 ESSLFTKEQVKDLPVGVHDVGLRFLFIDISVRNVCACINNEEDRVYVNVVDLKEVWAN 418
 QY 397 KEDRLRKAMEATITTELRPEARNRDSLETVPTTKLRSYADRIRASELEKCLKOVGED 456
 DB 419 KEDRLRKAMEAQSIITESSKQFEARNRDSLETVPTTKLRSYADRIRASELEKCLKOVGED 478
 QY 457 ALTKMRRARIELSTGVNKLHGLPLOHLRCDGSDRSLDETLNHLNRNMFSLDMEKA 516
 DB 479 -IPKTRRAVDDLSRGVNVKLHGLPLOHLRCDGSDRSLDETLNHLNRNMFSLDMEKA 537
 QY 517 IIEQKIKAKVEKTK 530
 DB 538 VLEQKIRAKVEQNK 551

RESULT 6

HMAL ARATH STANDARD; PRT; 543 AA.
 AC P42804; Q9LOB9;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Glutamyl-tRNA reductase 1, chloroplast (update)
 GN HEWA1 OR HEMA OR ATIG58290 OR F19C14.9
 OS Arabidopsis thaliana (Mouse-ear cress)
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsie.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia; TISSUE=Leaf;
 EX MEDLINE=94198598; PubMed=7808550;
 RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buchler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buchler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,

RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.G.,
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziani A.,
 RA Milgitscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 408:816-820(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
 RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
 RT SSP consortium (Salk/Stanford/PGEC).";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 CC -|- CATALYTIC ACTIVITY: Glutamyl-tRNA(Glu) + NADPH = glutamate-1-
 CC semialdehyde + NADP(+) + tRNA(Glu).
 CC -|- PATHWAY: Porphyrin biosynthesis by the C5 pathway; first step.
 CC Involved in chlorophyll biosynthesis.
 CC -|- SUBCELLULAR LOCATION: Chloroplast.
 CC -|- TISSUE SPECIFICITY: PRESENT IN ALL TISSUES TESTED. SLIGHTLY MORE
 CC ABUNDANT IN LEAVES.
 CC -|- INDUCTION: By light.
 CC -|- SIMILARITY: Belongs to the glutamyl-tRNA reductase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; U03774; AAA19118.1; -;
 DR EMBL; AC008051; AAF82258.1; -;
 DR EMBL; AY072223; AAL60044.1; -;
 DR PIR; E96616; E96616.
 DR HSSP; Q42843; 1B29.
 DR InterPro; IPR000343; Glutr.
 DR Pfam; PF00745; Glutr dimer; 1.
 DR Pfam; PF05201; Glutr N; 1.
 DR Pfam; PF05200; Glutr NAD bind; 1.
 DR TIGRFAMs; TIGR01035; hemA; 1.
 DR PROSITE; PS00747; Glutr; 1.
 KW Porphyrin biosynthesis; Chlorophyll biosynthesis; Oxidoreductase;
 KW NADP; Transit peptide; Chloroplast; Multigene family.
 FT TRANSIT 1 ?
 FT CHAIN ? 543
 FT DOMAIN 23 33 POLY-SER.
 FT DOMAIN 305 308 POLY-VAL.
 FT CONFLICT 62 62 E -> V (IN REF. 1).
 FT CONFLICT 169 169 K -> N (IN REF. 1).
 FT CONFLICT 291 291 K -> L (IN REF. 1).
 FT CONFLICT 542 542 Q -> H (IN REF. 1).
 SQ SEQUENCE 543 AA; 59515 MW; 04A095FEC96CC014 CRC64;
 Query Match 67.9%; Score 1770.5; DB 1; Length 543;
 Best Local Similarity 69.5%; Pred. No. 9e-90;
 Matches 360; Conservative 59; Mismatches 86; Indels 13; Gaps 4;
 QY 23 SSSALCORVAG-----GGRERSGVVRCDAAGVQAQAQAAKAAVALEQFKIS-A 72
 DB 28 SSSSVSOTPLGLNGVRVLPKNRTRRGLIQKARCELSASSDSNAASASISALEQKNSAA 87
 QY 73 DRYMKERSTIAVLGLSVHTAPVEMREKLAVALWELPRALIOELTSLNHIIEAAVLSTCNRM 132
 DB 88 DRYTKERSIIVIGLSIHTAPVEMREKLAIPAENWPRALIELCOLNHIIEAAVLSTCNRM 147


```

QY 133 EIYVVALSWNRGIREVVDWMSKSGIPASELREHLFILRSSDATRHLFEVSSAGLSVLG 192
DB 148 EIYVVALSQHRRGKREVTETMSKSGIPVSEICQHRFLYLNKDATQHI FEVSSAGLSVLG 207
QY 193 EQQILAQVQVVRSGNSGGLKNDIRPKDAITAGKRVSRSTNTSSGAVSVSSAAVELA 252
DB 208 EQQILAQVQVVRSGNSGGLKNDIRPKDAITAGKRVSRSTNTSSGAVSVSSAAVELA 267
QY 253 LMKLPKSEALSARMMLIGAGKMGKLVIKHLVAKGCKKVVVNRVERVDAIREEMKDIEI 312
DB 268 LMKLPQSSNVSAARMCVIGAGKMGKLVIKHLVAKGCKKVVVNRSEERVSAREEMPGEI 327
QY 313 VYRPLSDMYQAAAEADVFTSTASSETSLFAKEHAALPVSTMGGVRLFVDISVRNV 372
DB 328 IYRPLDEMLACASEADVFTSTASSETPLFKHEHVENLPQASPEVGGGLHFHFDISVRNV 387
QY 373 ACVSEVGARVYVNDLKEVVEANKEDRLRKAMEAQTIITEELRFRFEARWDSLETVP 432
DB 388 SCVGEVETARVYVNDLKEVVAANKEDRMKAMEAQTIITESTQFEARWDSLETVP 447
QY 433 KURSVADIRASELEKCLQKVGEDALTQKORRAIBELSTGIVNKLHGLPQHLRCDGSDS 492
DB 448 KLRVAERIRVAELEKCKSKMGDD-INKKTTRAVDDLSRGIVNRLHGLPQHLRCDGSDS 506
QY 493 RTLDETLENHNLNMFSLDMKALIEQKIKAKVEKTO 530
DB 507 RTLSETLENHNLNRMVGL--EKDILEKIKAMAEQQO 542

RESULT 7
HMA2 ARATH
ID HMA2 ARATH STANDARD; PRT; 530 AA.
AC P49234; O04950;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glutamyl-tRNA reductase 2, chloroplast precursor (BC 1.2.1.-) (Glutr).
GN HMA2 OR AT1G09940 OR F21M12.33.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsia.
ON NCBI_TaxID=3702;
RX SEQUENCE FROM N.A.
RC STRAIN=sv. Columbia;
RC MEDLINE=96189258; PubMed=8605295;
RA Kumar A.M., Ceankovszki G., Soell D.;
RT "A second and differentially expressed glutamyl-tRNA reductase gene
RL from Arabidopsis thaliana."
RN Plant Mol. Biol. 30:419-426 (1996).

SEQUENCE FROM N.A.
RC STRAIN=sv. Columbia;
RC MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egtu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis

```

```

RT thaliana."
RL Nature 408:816-820 (2000).
CC -! CATALYTIC ACTIVITY: Glutamyl-tRNA(Glu) + NADPH = glutamate-1-
CC semialdehyde + NADP(+) + tRNA(Glu).
CC -! PATHWAY: Porphyrin biosynthesis by the C5 pathway; first step.
CC Involved in chlorophyll biosynthesis.
CC -! SUBCELLULAR LOCATION: Chloroplast (By similarity).
CC -! TISSUE SPECIFICITY: EXPRESSED IN LOW LEVELS IN ROOTS AND FLOWERS.
CC -! SIMILARITY: Belongs to the glutamyl-tRNA reductase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U27118; AAB01674.1; -.
CC EMBL; AC000132; AAB00749.1; -.
CC PIR; G86233; G86233.
CC PIR; S65773; S65773.
CC HSSP; Q42843; 1B29.
CC InterPro; IPR000343; Glutr.
CC InterPro; IPR000594; Thif domain.
CC Pfam; PF00745; Glutr_dimer; 1.
CC Pfam; PF05201; Glutr_N; 1.
CC Pfam; PF05200; Glutr_NAD_bind; 1.
CC TIGRFAMs; TIGR01035; hemA; 1.
CC PROSITE; PS00747; GLUTR; 1.
CC Porphyrin biosynthesis; Chlorophyll biosynthesis; Oxidoreductase;
KW NADP; Transient peptide; Chloroplast; Multigene family.
FT TRANSIT 1 64 CHLOROPLAST (POTENTIAL).
FT CHAIN 65 530 GLUTAMYL-TRNA REDUCTASE 2.
FT CONFLICT 410 410 E -> D (IN REF. 2).
SQ SEQUENCE 530 AA; 58305 MW; 0CB225296BA3D1A8 CRC64;

Query Match 64.5%; Score 1680; DB 1; Length 530;
Best Local Similarity 67.2%; Pred. No. 7.6e-85;
Matches 350; Conservative 60; Mismatches 85; Indels 26; Gaps 9;

QY 20 PRGSSALCQVAG-----GRRRSVGV---RCDAGVEAQAAVAAVAALEQPKI 70
DB 23 PTYSSSPAPLDVIGIRALPMNRKRGIIQARCEISPSN-----KAASISALEQKT 75
QY 71 SA-DRYMKERSTIAVIGLSVHTAPVEMREKLAVALAEELMPRAIQELTSLNHIEAAVLSTC 129
DB 76 SAIDRYTKERSISIVVIGLSIHTAPVEMREKLAIPAEWPRALAEELCGLNHIEAAVLSTC 135
QY 130 NRMEIYVVALSWNRGIREVVDWMSKSGIPASELREHLFILRSSDATRHLFEVSSAGLSL 189
DB 136 NRMEIYVVALSQHRRGKREVTETMSKSGIPVSEICQHRFLYLNKDVTQHI FEVSSAGLSL 195
QY 190 VLGEQILAQVQVVRSGNSGGLKNDIRPKDAITAGKRVSRSTNTSSGAVSVSSAAV 249
DB 196 VLGEQILAQVQVVRSGNSGGLKNDIRPKDAITAGKRVSRSTNTSSGAVSVSSAAV 255
QY 250 ELALMKLPK-SEALSARMMLIGAGKMGKLVIKHLVAKGCKKVVVNRSEERVSAREEM- 307
DB 256 ELALMKLPKSESHASSARMMLVVGAGKMGKLVIKHLVAKGCTQWVVVNRSEERVAARNEMP 315
QY 308 KDIEIYRPLSDMYQAAAEADVFTSTASSETSLFAKEHAALPVSTMGGVRLFVDISV 367
DB 316 PGVEIYYKPLDEMLSCAAEADVFTSTASSETPLFKHEVETLPPVRD-----ARLFDVDSV 371
QY 368 PRNVSAVSEVGAARVYVNDLKEVVEANKEDRLRKAMEAQTIITEELRFRFEARWDSLET 427
DB 372 PRNVSCVAIEDGTFRVNDLKEVVAANKEDRMKAMEAQTIITESTQFEARWDSLET 431
QY 428 VPTIKLRSYADIRASELEKCLQKVGEDALTQKORRAIBELSTGIVNKLHGLPQHLR 487
DB 432 VPTIKLRSYADIRASELEKCLQKVGEDALTQKORRAIBELSTGIVNKLHGLPQHLR 490

```



```

DR InterPro; IPR000343; GlutR.
DR Pfam; PF00745; GlutR dimer; 1.
DR Pfam; PF05201; GlutR_N; 1.
DR Pfam; PF05200; GlutR_NAD bind; 1.
DR TIGRFAMs; TIGR01035; hemA; 1.
DR PROSITE; PS00747; GLUTR; 1.
KW Porphyrin biosynthesis; Chlorophyll biosynthesis; Oxidoreductase;
KW NADP; Complete proteome.
FT ACT_SITE 50 50
FT ACT_SITE 99 99
FT ACT_SITE 33 33
FT CONFLICT 33 33
FT CONFLICT 48 48
FT CONFLICT 174 174
FT CONFLICT 280 280
FT CONFLICT 374 380
SQ SEQUENCE 428 AA; 47909 MW; 06A470955F9D3C64 CRC64;

Query Match 34.9%; Score 909.5; DB 1; Length 428;
Best Local Similarity 47.0%; Pred. No. 6.6e-43;
Matches 204; Conservative 77; Mismatches 138; Indels 15; Gaps 7;

QY 82 IAVIGLSVHTAPVEMREKLAVAEELWPAIQELTSLNHIEEAVALSTCNRMIEIYVALSW 141
DB 3 IAVVGLSHKTAPEVEIREKLSIQEAKLEALTHLSYPHIEBVTIVISICNRLIYAVVTD 62
QY 142 NRGIREVVDWMSKSGIPASELREHLFILRSSDATRHLFEVSAGLDSLVLGEGQILAQVK 201
DB 63 DHGIREVTOFUSEHKLPHSLRQHLFVLLHEDAVHMRVAAGLDSLVLGEGQILAQVK 122
QY 202 QVVRSGQSGGLGKNDIRMFKDAITAGKVRSETNISSGAVSVSSAAVELALMKLPKSEA 261
DB 123 NTHKLGQQYNGIKTLNRLFKQALTAGKVRVETTSIGTGAVSISAAVELAQI---KAEN 179
QY 262 LSA-RMLLIGAGCKMGLVKKHVAKGCKKVVVNRVSRVDAIREMKDIEIVRPLSDM 320
DB 180 LAACRVTLIGAGCKMGLVKKHVAKGCKKVVVNRVSRVDAIREMKDIEIVRPLSDM 239
QY 321 YQAAEADVFTSTASSTSLFAKEAE-ALPVSVDTMGGVRLFDIVSRNVSAVCSVFG 379
DB 240 MTVIAESHLVFTSTSATPELDRLAKLEWVLAQNQLM----LF-DISVPRNVHTDVNBLV 294
QY 380 AARVYNVDLKEVPEANKEDRLRKAMEAQTIITEELRRFEAWRDSLETVPTIKKLSYAD 439
DB 295 NVQAFNVDDLKAVVAQNVESRRKMAQEAERLEIEAFDIWRSLETVSTISLSRSKIE 354
QY 440 RIRASELEKLOKVEDALTKWRRAIEELSTGIYKVLHGLHQLRCDGSDSRITLDEL 499
DB 355 TIREQELKALSRLGSE-FGDKHQEIVETALRGIVNKLHDPVQLRA---QQDVEARR 409
QY 500 ENMHALNRMFSLDM 513
DB 410 RCMQTLQWLFNLDV 423

RESULT 10
HEMI_SYNY3
ID -HEMI_SYNY3 STANDARD; PRT; 427 AA.
AC P28463;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glutamyl-tRNA reductase (EC 1.2.1.-) (GlutR).
GN HEMA OR SLR1808.
OS Synecocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
OX NCBI_TaxId=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92235044; PubMed=1569081;
RA Verkamp E., Jahn M., Jahn D., Kumar A.M., Soell D.;
RT "Glutamyl-tRNA reductase from Escherichia coli and Synecocystis
PT 6803: Gene structure and expression.";
RL J. Biol. Chem. 267:8275-8280(1992).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93093989; PubMed=1459859;
RA Grimm B.;
RT "Identification of a hemA gene from Synecocystis by complementation
FT of an E. coli hemA mutant";
RL Hereditas 117:195-197(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hikosawa M., Sugiura M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuura A., Muraki A., Nakazaki N., Naruo K.,
RA Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
RA Yamada M., Yasuda M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synecocystis sp. strain PCC6803. II. Sequence determination of the
entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -1- CATALYTIC ACTIVITY: Glutamyl-tRNA(Glu) + NADPH = glutamate-1-
semialdehyde + NADP(+) + tRNA(Glu).
CC -1- ENZYME REGULATION: FEEDBACK INHIBITION BY HEME.
CC -1- PATHWAY: Porphyrin biosynthesis by the C5 pathway; first step.
CC Involved in chlorophyll biosynthesis.
CC -1- SUBUNIT: Homooctamer.
CC -1- SIMILARITY: Belongs to the glutamyl-tRNA reductase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL; M84218; AAA27289.1; -.
DR EMBL; X65963; CAA46779.1; -.
DR EMBL; D90908; BAA17738.1; ALT_INIT.
DR HSSP; Q42843; 1B29.
DR HAWAP; MF_00087; -.
DR InterPro; IPR000343; GlutR.
DR InterPro; IPR000594; Thif domain.
DR Pfam; PF00745; GlutR_dimer; 1.
DR Pfam; PF05201; GlutR_N; 1.
DR Pfam; PF05200; GlutR_NAD bind; 1.
DR TIGRFAMs; TIGR01035; hemA; 1.
DR PROSITE; PS00747; GLUTR; 1.
KW Porphyrin biosynthesis; Chlorophyll biosynthesis; Oxidoreductase;
KW NADP; Complete proteome.
FT ACT_SITE 50 50
FT ACT_SITE 99 99
FT ACT_SITE 33 33
SQ SEQUENCE 427 AA; 47525 MW; 5F5D9D744A900D7D CRC64;

Query Match 34.7%; Score 904; DB 1; Length 427;
Best Local Similarity 44.7%; Pred. No. 1.3e-42;
Matches 194; Conservative 81; Mismatches 147; Indels 12; Gaps 4;

QY 82 IAVIGLSVHTAPVEMREKLAVAEELWPAIQELTSLNHIEEAVALSTCNRMIEIYVALSW 141
DB 3 IAVVGLSHKTAPEVEIREKLSIQEAKLEALTHLSYPHIEBVTIVISICNRLIYAVVTD 62
QY 142 NRGIREVVDWMSKSGIPASELREHLFILRSSDATRHLFEVSAGLDSLVLGEGQILAQVK 201
DB 63 EKGVEITQFLSETGNIPLATLRYLFTLLHEDAVHMRVAAGLDSLVLGEGQILAQVR 122
QY 202 QVVRSGQSGGLGKNDIRMFKDAITAGKVRSETNISSGAVSVSSAAVELALMKLPKSEA 261
DB 123 TTHKLGQKYKVGRLDLRLFKQALTAGKVRVETDTGTGAVSISAAVELVHVRQVLS-- 180
QY 262 LSARMLLIGAGCKMGLVKKHVAKGCKKVVVNRVSRVDAIREMKDIEIVRPLSDMY 321
DB 181 -SOKTVIIGAGCKMGLVKKHVAKGCKKVVVNRVSRVDAIREMKDIEIVRPLSDMY 239

```

```

QY 322 QAAEADVVFTSTASSETSLFAKEHAELPPVSDTMTGGVRLFYDVISVPRNVSACVSEVGAA 381
DB 240 TAAAGDVTSTGATEPILNCENLTGCVINKSL-----MLVDISVPRNVAADVHMEQV 295
QY 382 RYVNDVLLKEVVEANKEDRLKAMEAQTIITEELRRFEAWRDSLETPTIKKLSYADRI 441
DB 296 RAFNVDDDKVVAQNOARRQARQAEALLEEIEAFLDWMRSLETPTVPTISSRSKVEDI 355
QY 442 RASELEKICLOKVGEDALTCKMRRATEELSTGVNKLHGLHQLHLCDCGSDSRLLDETLEN 501
DB 356 REQELEKALSRLGSE-FAEKHGVIEALTRGIVNKLHHPVQLRA----QODIEARKQC 410
QY 502 MHALNRMFSLDMK 515
DB 411 LRSKMLPDEVEE 424

RESULT 11
HEMI SYNEL STANDARD; PRT; 426 AA.
AC Q8DI53;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glutamyl-tRNA reductase (EC 1.2.1.-) (GluTR).
GN HEMA OR TLL1738.
OS Synecococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=32046;
RN 11
RP SEQUENCE FROM N.A.
RX STRAIN=BP-1;
RX MEDLINE=2225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriyuchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shampo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1.";
RL DNA Res. 9:123-130(2002).
CC -|- CATALYTIC ACTIVITY: Glutamyl-tRNA(Glu) + NADPH = glutamate-1-
CC semialdehyde + NADP(+) + tRNA(Glu).
CC -|- PATHWAY: Porphyrin biosynthesis by the C5 pathway; first step.
CC Involved in chlorophyll biosynthesis.
CC -|- SIMILARITY: Belongs to the glutamyl-tRNA reductase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL; AF005375; BAC09290.1; -.
DR HAMAP; MF_00087; -.
DR InterPro; IPR00343; GluTR.
DR InterPro; IPR000594; Thif domain.
DR Pfam; PF00745; GluTR dimer; 1.
DR Pfam; PF05201; GluTR N; 1.
DR Pfam; PF05200; GluTR NAD_bind; 1.
DR TIGRFAMs; TIGR01035; hema; 1.
DR PROSITE; PS00747; GLUTR; FALSE NEG.
KW Porphyrin biosynthesis; Chlorophyll biosynthesis; Oxidoreductase;
KW NADP; Complete proteome.
FT ACT_SITE 50 50 NUCLEOPHILE (BY SIMILARITY).
FT ACT_SITE 99 99 BASE (BY SIMILARITY).
SQ SEQUENCE 426 AA; 47596 MW; D84CE5A1D2AA777E CRC64;

Query Match 34.3%; Score 893.5; DB 1; Length 426;
Best Local Similarity 44.2%; Pred. No. 4.9e-42;
Matches 193; Conservative 87; Mismatches 134; Indels 23; Gaps 5;

```

```

QY 82 IAVIGLSVHTAPVEMREKLAVAEELMPRAIOELTSLNHIEEAVALSTCNRMETVYVALSW 141
DB 3 IAVIGLSHTAPVDVREKLSVPEDYRERALQHLGCAHQEATILSTCNRLIYVTSOT 62
QY 142 NRGIREVVDWMSKSGIPASBLREHLFILRSSDATRHLFEVSAGLDSLVLGEGQILAQVK 201
DB 63 EVGVREVHQFLSEWHSIPLPOLRPLFYLLHQDAVMHLMRVASGLDSLVIGEGQILSQVK 122
QY 202 QVVRSGQSGGLKXNIDRMFKDAITAGKRVSETHISSGAVSVSSAAVELALMKLPKSEA 261
DB 133 RCHQGGQYKSGIPILNRVFTGATAGKRVSTETSGTGAIVSISAAVELADRLQNLQ- 181
QY 262 LSARMLLIGAGKMGKLVIKHLVAKCKKVVVNRVSRVDAIREEMKDIIEIVRPLSDMY 321
DB 182 -NCRIAVGAGKMSLVVQHLLIARVGVKIRIINRSLERAQELAQOQFPEVRFELFTWIDL 240
QY 322 QAAEADVVFTSTASSETSLFAKEHAELPPVSDTMTGGVRL-----FVDISVPRNVSACV 375
DB 241 PIVAMDLVFTSTAATEPLDDR-----DNLGAVLVGDSRLAIDISVPRNVHANV 290
QY 376 SEVGARVYVDDDLKEVVEANKEDRLKAMEAQTIITEELRRFEAWRDSLETPTIKKLR 435
DB 291 TELGTVQLFNVDDDLQAVVAQNOARRQAEAGILEBELETFLAWHHALETPTPIRSUR 350
QY 436 SYADRIRASELEKICLOKVGEDALTCKMRRATEELSTGVNKLHGLHQLHLCDCGSDSRLL 495
DB 351 QKMEAIRTOELEKALSRLGSE-FADKHGVIEAMTRTIIKILHDPVQLQ-----SORDL 405
QY 496 DETLENMHALNRMFSLD 512
DB 406 ESRQRAMQTLQDLFNLE 422

RESULT 12
HEMI CYAPA
ID HEMI CYAPA STANDARD; PRT; 432 AA.
AC P48077;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glutamyl-tRNA reductase (EC 1.2.1.-) (GluTR).
GN HEMA.
OS Cyanophora paradoxa.
OC Eukaryota; Glaucocystophyceae; Cyanophoraceae; Cyanophora.
OX NCBI_TaxID=2762;
RN 11
RP SEQUENCE FROM N.A.
RX STRAIN=UTEX LB 555 / Pringsheim;
RX Stirewalt V.L., Michalowski C.B., Loeffelhardt W., Bohnert H.J.,
RX Bryant D.A.;
RT "Nucleotide sequence of the cyanelle DNA from Cyanophora paradoxa.";
RT Plant Mol. Biol. Rep. 13:327-332(1995).
RN 12
RP SEQUENCE FROM N.A.
RX STRAIN=UTEX LB 555 / Pringsheim;
RX Loeffelhardt W., Stirewalt V.L., Michalowski C.B., Annarella M.,
RX Farley J.Y., Schlachter W.M., Chung S., Newmann-Spallart C.,
RX Steiner J.M., Jakowitsch J., Bohnert H.J., Bryant D.A.;
RT "The complete sequence of the cyanelle genome of Cyanophora paradoxa:
RT the genetic complexity of a primitive plastid.";
RL (In) Schenk H.E.A., Herrmann R., Jeon K.W., Mueller N.E.,
RL Schwemmler W. (eds.);
RL Eukaryotism and Symbiosis, pp.40-48, Springer-Verlag, Heidelberg
RL (1997).
RN 13
RP SEQUENCE OF 407-432 FROM N.A.
RX STRAIN=UTEX LB 555 / Pringsheim;
RX MEDLINE=90092562; PubMed=2403527;
RA Bryant D.A., Stirewalt V.L.;
RT "The cyanelle genome of Cyanophora paradoxa encodes ribosomal
RT proteins not encoded by the chloroplast genomes of higher plants.";
RL FEBS Lett. 259:273-280(1990).

```

```

CC -!- CATALYTIC ACTIVITY: Glutamyl-tRNA (Glu) + NADPH = glutamate-1-
CC semialdehyde + NADP(+) + rRNA (Glu).
CC -!- PATHWAY: Porphyrin biosynthesis by the C5 pathway; first step.
CC Involved in chlorophyll biosynthesis.
CC -!- SIMILARITY: Belongs to the glutamyl-tRNA reductase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U30821; AAA81194.1; -.
DR EMBL; X17063; -. NOT_ANNOTATED_CDS.
DR FIR; T06851; T06851.
DR HSSP; Q42843; 1B29.
DR HAMAP; MF_00087; -. 1.
DR InterPro; IPR000343; GlutR.
DR Pfam; PF00745; GlutR dimer; 1.
DR Pfam; PF05201; GlutR_N; 1.
DR Pfam; PF05200; GlutR_NAD_bind; 1.
DR TIGRfams; TIGR01035; hema; 1.
DR PROSITE; PS00747; GLUTR; FALSE NEG.
KW Porphyrin biosynthesis; Chlorophyll biosynthesis; Oxidoreductase;
KW NADP; Cyanelle.
FT ACT_SITE 50 50 NUCLEOPHILE (BY SIMILARITY).
FT ACT_SITE 99 99 BASE (BY SIMILARITY).
FT ACT_SITE 99 99 BASE (BY SIMILARITY).
SQ SEQUENCE 432 AA; 48542 MW; F14378FC72FFD05B CRC64;

Query Match 33.5%; Score 872.5; DB 1; Length 432;
Best Local Similarity 43.5%; Pred. No. 7e-41;
Matches 189; Conservative 86; Mismatches 148; Indels 11; Gaps 4;

QY 82 IAVIGLSVHTAPVEMREKLAVALAEELWPRALQELTSLNHEEAVALSTCNRMIEIYVALSW 141
DB 3 IIVVGLSHKTPADVDFREKLSIPKVRIGEAIRLNCYPHIEEVAIILSTCNRLIEIYLTSDT 62
QY 142 NRGIREVVDWMSKSGIPASSELREHLFILRSSDATRHLFEVSAGLDSLVLGEGQILAQV 201
DB 63 YQGIREATQPLADSDLSLELURQHLFILLHQDAVMHLMRVTAGLDSLIIIGEGQILSQVK 122
QY 202 QVVRSGQSGGLGNIDRMFKDAITAGKVRSETNISSGAVSVSSAAVELALMLKPKSEA 261
DB 123 QCYQLGQQYQIGPVNLNIFQAISAGKVRVETQISTGAVSISSAAVELAQIK--KQDL 180
QY 262 LSARMLLIGAGKMGKLVILKLVAKCKKVVVNVRSVERVDAREMKDIEIVYRPLSDMY 321
DB 181 RTANITILGAGMGRLVVOHLLSKRVKDVINVRSVERAKLLVDQFKEANINIYNSLSELK 240
QY 322 QAAAEADVFTSTASSETSLFAKEHAELPPVSDTMGGVRLFVDISVPRNVSAVCSEVGAA 381
DB 241 TILQNSDIVFTGTSQEPITPE-----LINCDELPSLEMLFDIAPRNVDPNVSQFDNI 296
QY 382 RVYVNDLKEVVEANKEDRLRKAMEAQTIITEELRFEAWRDSLETVPTIKKLSYADRI 441
DB 297 KVFNVVDLLKVVVSNQOQTRKWKAAAEILLEELSAFNWWSLEAFTINKLEKAEII 356
QY 442 RASLEKCLKVGEDALTKKRRRAIEELSTGVNKLHGPIQLHRCGSDSRITDETLEN 501
DB 357 RVKELEKAIKSLRGNEFVSDH-QEIVESITRGIVNKLHDPVQLRA-----QODIEIRGA 411
QY 502 MHALNRMFSLDMEK 515
DB 412 LKILQTLFNLDTIK 425

RESULT 13
HEM1 VIBVU
ID HEM1 VIBVU
AC Q8DF8;
DT 15-SEP-2003 (Rel. 42, Created)

```

```

DT 15-SEP-2003 (Rel. 42, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glutamyl-tRNA reductase (EC 1.2.1.1-) (Glutr).
GN HEMA OR VV10254.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: Glutamyl-tRNA (Glu) + NADPH = glutamate-1-
CC semialdehyde + NADP(+) + rRNA (Glu).
CC -!- PATHWAY: Porphyrin biosynthesis by the C5 pathway; first step.
CC -!- SIMILARITY: Belongs to the glutamyl-tRNA reductase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE016797; AA008790.1; -.
DR HAMAP; MF_00087; -. 1.
DR InterPro; IPR000343; GlutR.
DR Pfam; PF00745; GlutR dimer; 1.
DR Pfam; PF05201; GlutR_N; 1.
DR Pfam; PF05200; GlutR_NAD_bind; 1.
DR TIGRfams; TIGR01035; hema; 1.
DR PROSITE; PS00747; GLUTR; 1.
KW Porphyrin biosynthesis; Oxidoreductase; NADP; Complete proteome.
FT ACT_SITE 50 50 NUCLEOPHILE (BY SIMILARITY).
FT ACT_SITE 97 97 BASE (BY SIMILARITY).
FT ACT_SITE 97 97 BASE (BY SIMILARITY).
SQ SEQUENCE 419 AA; 46057 MW; 834113C720656CBD CRC64;

Query Match 24.8%; Score 646; DB 1; Length 419;
Best Local Similarity 36.6%; Pred. No. 1.5e-28;
Matches 154; Conservative 81; Mismatches 144; Indels 42; Gaps 9;

QY 81 TIAVIGLSVHTAPVEMREKLAVALAEELWPRALQELTSLNHEEAVALSTCNRMIEIYVALS 140
DB 2 SLAIGINHHTASVDLREKVAFGPKLANALQQLSQHEAVNGSVILSTCNRTVEYCDVKS 61
QY 141 WNRGIREVVDWMSKSGIPASSELREHLFILRSSDATRHLFEVSAGLDSLVLGEGQILAQV 200
DB 62 GARS--KVIDWLSQFHQIGLEELKPSLYVVEEQAAIRHLMRVSCGLDSLVLGEPQILGQV 119
QY 201 QVVRSGQSGGLGNIDRMFKDAITAGKVRSETNISSGAVSVSSAAVELA---LMLKP 257
DB 120 QKAYSDESQOQVNDASLEKLFQKTFVAKVRVETDIGNRVSVAYAACTLAKHIPSELE 179
QY 258 KSEALSARMLLIGAGKMGKLVILKLVAKCKKVVVNVRSVERVDAREMKDIEIVYRPL 317
DB 180 KS-----TVLVGAGETIELVAKHLHANGCKSMIVANTRERALTAEQF-DAQVI--SL 231
QY 318 SDMYQAAAEADVFTSTAS-----ETSLFAKEHAELPPVSDTMGGVRLFVDISV 367
DB 232 QEIPNHLAKADIVISSTASPLPIIGKMWETALKQRHQP-----MLLVDAIV 279
QY 368 PRNVSAVCSEVGAARVYVNDLKEVVEANKEDRLRKAMEAQTIITEELRFEAWRDSLET 427
DB 280 PRDVEAQVQGLNDAYLYTVDDLSQSIDSNIEQKRVKVEAIQAEIQAIEVAESAFAFMSWLSQA 339
QY 428 VPTIKKLSYADRIASELEKCLKQV--GEDALTKKRRRAIEELSTGVNKLHGPIQLH 485
DB 340 VDSIREYQSANETRELDLSKALLSLESGSDP-----EKLRLSNRLTNKLIHAPTRAL 394

```

```

QY 486 R 486
Db 395 Q 395

RESULT 14
HEMI_HELMO
ID HEMI_HELMO STANDARD; PRT; 443 AA.
AC Q9ZG6;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glutamy1-tRNA reductase (EC 1.2.1.-) (Glutr).
GN HEMA.
OS Hellobacillus mobilis.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Hellobacteriaceae;
OC Hellobacillus.
OX NCBI_TaxID=28064;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99061957; PubMed=9843979;
RA Xiong J., Inoue K., Bauer C.E.;
RT "Tracking molecular evolution of photosynthesis by characterization of
a major photosynthesis gene cluster from Hellobacillus mobilis.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:14851-14856(1998).
CC -!- CATALYTIC ACTIVITY: Glutamyl-tRNA(Glu) + NADPH = glutamate-1-
semialdehyde + NADP(+) + tRNA(Glu).
CC -!- PATHWAY: Porphyrin biosynthesis by the C5 pathway; first step.
CC -!- SIMILARITY: Belongs to the glutamyl-tRNA reductase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF080002; AAC84013.1; -
DR PIR; T31441; T31441.
DR HSP; Q42843; 1B29.
DR HAMAP; MF_00087; -; 1.
DR InterPro; IPR000343; Glutr.
DR Pfam; PF00745; Glutr dimer; 1.
DR Pfam; PF05201; Glutr N; 1.
DR Pfam; PF05200; Glutr NAD bind; 1.
DR TIGRfam; TIGR01035; hemA; 1.
DR PROSITE; PS00747; GLUTR; 1.
KW Porphyrin biosynthesis; Oxidoreductase; NADP.
FT ACT SITE 50 NUCLEOPHILE (BY SIMILARITY).
FT ACT SITE 99 BASE (BY SIMILARITY).
SQ SEQUENCE 443 AA; 49577 MW; E95C30E251A0C5F6 CRC64;

Query Match 24.7%; Score 643.5; DB 1; Length 443;
Best Local Similarity 34.4%; Pred. No. 2.3e-28;
Matches 157; Conservative 98; Mismatches 177; Indels 25; Gaps 9;

QY 82 IAVIGLSVHTAPVEMREKLAVALAEELWPRAIOELTSLNHEEAAVLSTCNMEIVVVALSW 141
Db 3 IFVVLGNHKSAPVEVREKLSFTEAQLSEALHKLQMGAGIEGCGILSTCNRTIYGASTDM 62
QY 142 NRGIREVVDWMSKSGIPASELREHLFILRSGDATRHLFEVSAGLDSLVLGEGQILAQVK 201
Db 63 EKGMTAVKRFVLEMGQLQPDQSKFYFVYHTLYDAIRHLFRVASGLDSNVLGETQLQGV 122
QY 202 QVRSQSGGLGKKNIDRMFKDAITAGKVRSETNISSGAVSSAAVELALMKLPKSEA 261
Db 123 TAYORSNCNDCSNGIVNTWFOQAITVGRVTRTETGIDQHPVSIYSTAVELAEQVLGLKG 182
QY 262 LSARMLLIGAGQMGKLVTKHLVAKGCKVWVNVRSVRVDAIREMKDIEIVYRPLSPMY 321
Db 183 RTA--MVLGAGMSVLTUKHLVAEGVDKIIIANRSVEKAEEUAKSCGGEATSF---ADVN 237

```

```

QY 322 QAAAEADVVTSTASSTSLFAKEHAELPPVSDTWGGVRL--FVDISVPENVSACVSEVGA 380
Db 238 HRLSEADILISCTAAHTVIRKSMVE---QVMDRRGGKVPFFDIAVRDIDPEVAQVPG 294
QY 381 ARVYNVDLKEVVEANKEDRLRKAMEAQTIITEELRRFEAMRDSLETPTIKKLSYADR 440
Db 295 THLYDIDAMQHVDRNLAEERRKCAAEABIIIEHEINQFWRWLSLFPVIFTIVGLKNGNQ 354
QY 441 IRASELEKCLKQVGEDALTKQRRRAIEBLSTQIVNKLHGLPQHLR--CDGSDSRTLDET 498
Db 355 IKEKELDRALCKLKH--LSEKEKLVGSLASSIVNQLLHDPITQLRHYAASPEGHLYSEI 412
QY 499 LEMHALNRMFSLDW----EKALIEQIKAKVETKTON 531
Db 413 LQN-----LFCLDVPQGRQKRVVH--YPAVEQRQN 441

RESULT 15
HEMI_NEIMB
ID HEMI_NEIMB STANDARD; PRT; 415 AA.
AC P56994;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glutamyl-tRNA reductase (EC 1.2.1.-) (Glutr).
GN HEMA OR NMB0576.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAINS=WC58 / Serogroup B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.P., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Citterone H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizzi M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
MC58.";
RL Science 287:1809-1815(2000).
CC -!- CATALYTIC ACTIVITY: Glutamyl-tRNA(Glu) + NADPH = glutamate-1-
semialdehyde + NADP(+) + tRNA(Glu).
CC -!- PATHWAY: Porphyrin biosynthesis by the C5 pathway; first step.
CC -!- SIMILARITY: Belongs to the glutamyl-tRNA reductase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB002413; AAF41004.1; -
DR PIR; C81183; C81183.
DR HSP; Q42843; 1B29.
DR TIGR; NMB0576; -; 1.
DR HAMAP; MF_00087; -; 1.
DR InterPro; IPR000343; Glutr.
DR Pfam; PF00745; Glutr dimer; 1.
DR Pfam; PF05201; Glutr N; 1.
DR Pfam; PF05200; Glutr NAD bind; 1.
DR TIGRfam; TIGR01035; hemA; 1.
DR PROSITE; PS00747; GLUTR; 1.
KW Porphyrin biosynthesis; Oxidoreductase; NADP; Complete proteome.
FT ACT SITE 50 NUCLEOPHILE (BY SIMILARITY).
FT ACT SITE 94 BASE (BY SIMILARITY).
FT ACT SITE 94 BASE (BY SIMILARITY).
SQ SEQUENCE 415 AA; 45480 MW; 6C5DD4A0D6B03894 CRC64;

```

```
Query Match      24.5%; Score 637.5; DB 1; Length 415;
Best Local Similarity 34.3%; Pred. No. 4.4e-28;
Matches 149; Conservative 96; Mismatches 163; Indels 27; Gaps 9;

QY 82 IAVIGLSVHTAPVEMREKLAVALPRAIQELTSLNHIEEAVALSTCNRMIEIYVVALSW 141
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
3 LTAVALGNHQTPALSTIREKLAFAAALPKAVRNLAARNAATEAVILSTCNRTLYCVGDS- 61

QY 142 NRGIREVDWMSKSGIPASELREHLFILRSSDATRHLFEVSAGLDSLVLGEGQILAQVK 201
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
62 ----EIIIRMLADYHSLPTEIIRPYLYALDMQETVRHAFRVACGLDSMWLGEPLIGQIK 117

QY 202 QVVRSGNSGGKGNIDRMFKDAITAGKRVRSSTNISSGAVSVSSAAVELALMKLPKSEA 261
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
118 DAVRVAQOEESGKKLNALFQKTFVAKVTRDTAVGENSVSASVVKLAEOIIPDIDG 177

QY 262 LSARMLLIGAGKXGKLVIKHLVAKGCKVYVNRSVRVDAIREEMKD-IEIVYRP--LS 318
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
178 LN--VLFIGAGEMIELVATYFAAKSPRLMTVANRTLARA---QELCDKLGYNABECLLS 231

QY 319 DMVQAAAADVVFTSTASSETSLFAKEHAALPPVSDTMGGVRLF-VDISVPRNVACVSE 377
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
232 DLPAILHDYDVVSVSTASQPIVGKGMVERALKQOSM---PLFMLDLAVPRDIEAEVGD 288

QY 378 VGAARVYVNDLKEVVVEANKEDRLRKAMEAQTIITEELRRFEAWRDSLETPTIKKLSY 437
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
289 LNDAYLYTVDDVNIYVQSGKEARQKAAAAAETLVSEKVAEFVRQQGROSQVPLIKALRDE 348

QY 438 ADRIASELEKCLQKVGEDALTKMMRAIEELSTGIVNKLHGPLQHLCDSGSDSRTLDE 497
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
349 GEKARKQVLENAMKQAKGATAEV---LERLSVQLTNKLLHSPQTTLNKAGEEDKDL-- 403

QY 498 TLENMHALNRMFSLD 512
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
404 ----VHAVAQIYHLD 414
```

Search completed: December 15, 2003, 12:55:14
Job time : 18 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 12:53:43 ; Search time 21 Seconds

(without alignments)

2431.694 Million cell updates/sec

Title: US-10-018-902-4

Perfect score: 2606

Sequence: 1 MATTSATTAATAAAATAKPP.....DMEKATTEQKIKAVKENTON 531

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_76:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2313	88.8	528	2 T05732	probable glutamyl-
2	2225	85.4	535	2 T04402	probable glutamyl-
3	2113	81.1	465	2 T05734	glutamyl-tRNA redu
4	1784.5	68.5	552	2 T10186	glutamyl-tRNA redu
5	1770.5	67.9	543	2 E96616	hypothetical prote
6	1680	64.5	530	2 S65773	glutamyl-tRNA redu
7	1677	64.4	530	2 G86233	hypothetical prote
8	1573	60.4	542	2 T10245	glutamyl-tRNA redu
9	1334	51.2	524	2 D84718	probable glutamyl
10	910.5	34.9	463	2 S77180	glutamyl-tRNA redu
11	909.5	34.9	428	2 AG1936	glutamyl-tRNA redu
12	872.5	33.5	432	2 T06851	glutamyl-tRNA redu
13	643.5	24.7	443	2 T31441	probable glutamyl
14	637.5	24.5	415	2 C81183	glutamyl-tRNA redu
15	628.5	24.1	415	2 A81920	glutamyl-tRNA redu
16	620	23.8	419	2 C82109	glutamyl-tRNA redu
17	616.5	23.7	459	2 H84030	glutamyl-tRNA redu
18	577.5	22.2	420	2 AB0246	glutamyl-tRNA redu
19	573.5	22.0	418	1 BVECHA	glutamyl-tRNA redu
20	573.5	22.0	418	2 C90843	glutamyl-tRNA redu
21	573	22.0	455	1 A35252	glutamyl-tRNA redu
22	572.5	22.0	418	2 H85700	glutamyl-tRNA redu
23	569.5	21.9	418	1 EVESHA	hypothetical prote
24	569.5	21.9	418	2 AC0720	glutamyl-tRNA redu
25	548.5	21.0	422	2 C83063	glutamyl-tRNA redu
26	544.5	20.9	432	2 A82533	glutamyl-tRNA redu
27	538.5	20.7	435	2 AG1431	glutamyl-tRNA redu
28	533	20.5	406	2 G70410	glutamyl-tRNA redu
29	523.5	20.1	448	2 F89950	glutamyl-tRNA redu

30 506 19.4 435 2 A81269 glutamyl-tRNA redu
31 505.5 19.4 350 2 F75259 glutamyl-tRNA redu
32 459 17.6 468 2 H70746 probable hema prot
33 450 17.3 392 2 H64317 glutamyl-tRNA redu
34 446.5 17.1 398 1 S51136 glutamyl-tRNA redu
35 444.5 17.1 437 2 F69496 glutamyl-tRNA redu
36 435.5 16.7 402 2 A69002 glutamyl-tRNA redu
37 430 16.5 404 2 T45026 glutamyl-tRNA redu
38 428.5 16.4 449 2 G64549 glutamyl-tRNA redu
39 423.5 16.3 436 2 B84329 glutamyl-tRNA redu
40 419.5 16.1 581 2 T36267 probable glutamyl-
41 417 16.0 467 2 S72907 glutamyl-tRNA redu
42 415.5 15.9 449 2 C71959 glutamyl-tRNA redu
43 400 15.3 432 2 F81400 glutamyl-tRNA redu
44 385 14.8 416 2 D72456 probable glutamyl-
45 369 14.2 400 2 T43856 glutamyl-tRNA redu

ALIGNMENTS

RESULT 1

T05732

probable glutamyl-tRNA reductase (EC 1.2.1.1.-) 1 precursor, chloroplast - barley
C:Species: Hordeum vulgare (barley)

C:Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 16-Jul-1999

C:Accession: T05732

R:Bougri, O.; Grimm, B.

Plant J. 9, 867-878, 1996

A:Title: Members of a low-copy number gene family encoding glutamyl-tRNA reductase are
A:Reference number: Z15440; MUID:96267576; PMID:8696365

A:Accession: T05732

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-528 <BOU>

A:Cross-references: EMBL:X86101; NID:g1039331; PIDN:CAA60054.1; PID:g1039332

A:Experimental source: cv. Klages

C:Genetics:

A:Gene: hema1

A:Map position: 5

A:Genome: nuclear

C:Superfamily: Glutamyl-tRNA reductase

C:Keywords: aminolevulinic biosynthesis; chloroplast; NADP; oxidoreductase; porphyrin
F:1-44/Domain: transit peptide (chloroplast) #status predicted <NTP>
F:45-528/Product: Glutamyl-tRNA reductase 1 #status predicted <NAT>

Query Match 88.8%; Score 2313; DB 2; Length 528;
Best Local Similarity 88.1%; Pred. No. 8.6e-122;
Matches 468; Conservative 27; Mismatches 32; Indels 4; Gaps 2;

Qy 1 MATTSATTAATAAAATAKPPRGSSALCORVAGGRRRRSGVVRCDAAAGVEAQAQAAKAA 60

Db 2 MAGATSA-TAAAGAFAAAKARGPAAACPMVLVAGGRRRRSGVVRCDAGG---DAQAASKAA 57

Qy 61 SVAALQFKISADRYMKERTIAVIGLSVHTAPVREKLAVALAEELWPAIQLTSLNHI 120

Db 58 SITALEQFKISADRYMKERKSSIAVIGLSVHTAPVREKLAVALAEELWPAIQLTSLNHI 117

Qy 121 EEAALSTCNRMIEIYVVALSWNRGIREVDWMKSGIPASELREHLFILRSSDATRHLP 180

Db 118 EEAALSTCNRMIEIYVVALSWNRGIREVDWMKSGIPASELREHLFILRSSDATRHLP 177

Qy 181 EYSAGLDSLVLGSGQILAQVKVVRSGQNSGGIGKQINDRMFKDAITAGKVRSTNLSG 240

Db 178 EYSAGLDSLVLGSGQILAQVKVVRSGQNSGGIGKQINDRMFKDAITAGKVRSTNLSG 237

Qy 241 AVSVSSAAVELALMKPKSEALSARMMLTGAGDKGKLVIGLVAKGCKKVVVNRVSERV 300

Db 238 AVSVSSAAVELAMKLPKSEALSARMMLTGAGDKGKLVIGLVAKGCKKVVVNRVSERV 297

Qy 301 DAIREMKDIEIYVRPLSDMYQAAAEADVFTTSTSETSLPAKEHAELPPVSDTGGVVR 360

Db 298 DAIREMKDIEIYVRPLTEYEAADADVFTTSTSETSLFTKEHAELPPISLAMGGVVR 357

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzili
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A>Title: Sequence and analysis of chromosomes 1 of the plant Arabidopsis.
A;Reference number: AB6141; PMID:21016719; PMID:11130712

A;Accession: E96616
A>Status: preliminary
A:Molecule type: DNA
A;Residues: 1-543 <STO>
A;Cross-references: GB:AB005173; NID:g8979944; PIDN:AAP82258.1; GSPDB:GN00141
C;Genetics:
A;Gene: F19C14.9
A;Map position: 1
C;Superfamily: glutamyl-crRNA reductase

Query Match 67.9%; Score 1770.5; DB 2; Length 543;
Best Local Similarity 69.5%; Pred.No.1.7e-91;
Matches 360; Conservative 59; Mismatches 86; Indels 13; Gaps 4;

Qy 23 SSSSLCORVAG-----GRRRSVVRCDAAGVEAOAAOVAKASVAALAEOKFIS-A 72
 :|::||
Db 28 SSSSVSQTPLEGLGVVLPKNNRTRGIIQKARCELSASSDSNASNAASfSALBOLKNAA 87
 :|::||

Qy 73 DRYMKERSTIAIGLVSHTPAVMREKLAVABELMPRAIOELTSLNHIEEAAVLSTCNRM 132
 :|::||
Db 88 DRYTKERSIIVIGLSIHITAPMEREKLAIPAEWPRAETAECLGANHIEEAVALSTCNRM 147
 :|::||

Qy 133 EIYVVALSWNRGIREVVDWKKSGIPASELRHUFLRSSDATRHLPFEVSAGLDLSVLG 192
 :|::||
Db 148 EIYVVALSOHRGVKEVTWNKSITGPVSEICQHRFLYNKOATQHIFEVSAGLDLSVLG 207
 :|::||

Qy 193 EGQLAQVKQVVRSSONGSGLGKNDIRMFKDITAGKVRSSETNISGGVASVSAAYELA 252
 :|::||
Db 208 EQLLAQVKQVKGCGVGNGFGRNISGLFKHAITYGKRVRTETNIASGASVSSAAVELA 267
 :|::||

Qy 253 LMCLKPESALSARMILLIGAGKMCKLVIKHLVAKGCCKVNVVNRSRVDAIREMCKDEI 312
 :|::||
Db 268 LMKLPOSSNSVARMCVIGAGKMCKLVIKHLMAGCTKVNVNRSERVSAIREMPGIEI 327
 :|::||

Qy 313 VTPLSDMYQAABAADVFTTSASTSIAPAHEBALPPVSDTMGGRVLFVDISVPNNVS 372
 :|::||
Db 328 IYRPDLMLACAAEDVFTSTASTPTPLFKHEVENLPQASPVEVGLRFHFVDISVPNNVG 387
 :|::||

Qy 373 ACYSDEVGAARYVDDLKEVYEANKEDRLRKAMEACTIITEELLRRPEARWRDSLTVPTIK 432
 :|::||
Db 388 SCGEVETARYVDDDKBVYAANKEDBRKAMEACTIITESQTPEAWRDLSLTVPTEIK 447
 :|::||

Qy 433 KLRSYADRIASELEKLCLOKVEDALTKMRRAIBELSITGINKLHGPHOHURCDGSDDS 492
 :|::||
Db 448 KRAYAESIRVAELEKCMKWGD--INKTYTRAVDDELKGI VNRLFPGHPHOHURCDGSDDS 506
 :|::||

Qy 493 RTLDTELNMHALNFMPSLDMEKAIIEOKIKAKVEKTQ 530
 :|::||
Db 507 RTLSETLENMALNRYMG|--EKDIILEBKAKAMAEQQQ 542
 :|::||

RESULT 6
S65773
glutamyl-t-RNA reductase (EC 1.2.1.-) precursor, chloroplast - Arabidopsis thaliana
C;Species: Arabidopsia thailana (mouse-ear cress)
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 16-Jul-1999
C;Accession: S65773
R;Kurata, A.M.; Ceankovszki, G.; Soell, D.
Plant Mol. Biol. 30, 419-426, 1996
A>Title: A second and differentially expressed glutamyl-t-RNA reductase gene from Arab
A;Reference number: S65773; PMID:96189258; PMID:8605295
A;Accession: S65773
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA
A;Residues: 1-530 <KUM>
A;Cross-references: EMBL:U27118; NID:g1049056; PIDN:AAB01674.1; PID:g1049057
A>Note: mRNA was also sequenced

C:Genetics:
A:Gene: HEMA2
A:Genome: nuclear
A:Introns: 80/1; 160/1
C:Superfamily: glutamyl-tRNA reductase
C:Keywords: aminolevulinate biosynthesis; chloroplast; NADP; oxidoreductase; porphyrin b
F:1-64/Domain: transit peptide (chloroplast) #status predicted <TNP>
F:65-530/Product: glutamyl-tRNA reductase 2 #status predicted <TNP>

Query Match 64.5%; Score 1680; DB 2; Length 530;
Best Local Similarity 67.2%; Pred. No. 1.8e-86;
Matches 350; Conservative 60; Mismatches 85; Indels 26; Gaps 9;

QY 20 PRGSSALCORVAG-----GRRRSVV---RCDAGVEAQAQAAVAAAEQFKI 70
DB 23 PTYSSPAPLDVIGIRALPMNNRKGIIQARCEISPSN-----KAASISALEQLKT 75
QY 71 SA-DRYMKERSTIAVIGLSVHTAPVEMREKLAVALAEELWPRALQELTSLNHIEEAVLSTC 129
DB 76 SAIDRYTKERSIVVIGLSIHTAPVEMREKLAIPAEWPRALAEELCGLNHIEEAVLSTC 135
QY 130 NMIEIYVALSNRGIREVVDWMSKSGIPASELREHLFILRSSDATRHLFEVSAGLDSL 189
DB 136 NMIEIYVALSNRGIREVVDWMSKSGIPASELREHLFILRSSDATRHLFEVSAGLDSL 195
QY 190 VLGEQIILAQVQVRSQNSGGLGNIDRMFKDAITAGKRVRSSETNISGAVSVSSAAV 249
DB 196 VLGEQIILAQVQVRSQNSGGLGNIDRMFKDAITAGKRVRSSETNISGAVSVSSAAV 255
QY 250 ELALMKLPK-SEALSARMLLIGAGKMGKLVIKHLVAKGCKVNVNRSVERVDAREEM- 307
DB 256 ELALMKLPK-SEALSARMLLIGAGKMGKLVIKHLVAKGCKVNVNRSVERVDAREEM- 315
QY 308 KDIEIVRPLSDMYQAAAEADVFTSTASETSLFAKEHAELPPVSDTMGGVRLFPVDISV 367
DB 316 PGVEIYKPLDEMLSCAAEADVFTSTASETSLFAKEHAELPPVSDTMGGVRLFPVDISV 371
QY 368 PRNVSCVSEVGAARVYVVDLKEVVEANKEDRLRKAMEAQTIIIEELRRFEAWRDSLET 427
DB 372 PRNVSCVSEVGAARVYVVDLKEVVEANKEDRLRKAMEAQTIIIEELRRFEAWRDSLET 431
QY 428 VPTIKLRSYADRIRASELEKLOKVGEDALTKQMRRAIEELSTGIVNKLHGPLOHLRC 487
DB 432 VPTIKLRSYADRIRASELEKLOKVGEDALTKQMRRAIEELSTGIVNKLHGPLOHLRC 490

RESULT 7
G86233
hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 24-Aug-2001
R:Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huijzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luross, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: AB6141; MUID:21016719; PMID:11130712
A:Accession: G86233
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-530 <STO>
A:Cross-references: GB:AE005172; NID:G2160186; PIDN:AA60749.1; GSPDB:GN00141
C:Genetics:

A:Map position: 1
C:Superfamily: glutamyl-tRNA reductase

Query Match 64.4%; Score 1677; DB 2; Length 530;
Best Local Similarity 67.0%; Pred. No. 2.6e-86;
Matches 349; Conservative 61; Mismatches 85; Indels 26; Gaps 9;

QY 20 PRGSSALCORVAG-----GRRRSVV---RCDAGVEAQAQAAVAAAEQFKI 70
DB 23 PTYSSPAPLDVIGIRALPMNNRKGIIQARCEISPSN-----KAASISALEQLKT 75
QY 71 SA-DRYMKERSTIAVIGLSVHTAPVEMREKLAVALAEELWPRALQELTSLNHIEEAVLSTC 129
DB 76 SAIDRYTKERSIVVIGLSIHTAPVEMREKLAIPAEWPRALAEELCGLNHIEEAVLSTC 135
QY 130 NMIEIYVALSNRGIREVVDWMSKSGIPASELREHLFILRSSDATRHLFEVSAGLDSL 189
DB 136 NMIEIYVALSNRGIREVVDWMSKSGIPASELREHLFILRSSDATRHLFEVSAGLDSL 195
QY 190 VLGEQIILAQVQVRSQNSGGLGNIDRMFKDAITAGKRVRSSETNISGAVSVSSAAV 249
DB 196 VLGEQIILAQVQVRSQNSGGLGNIDRMFKDAITAGKRVRSSETNISGAVSVSSAAV 255
QY 250 ELALMKLPK-SEALSARMLLIGAGKMGKLVIKHLVAKGCKVNVNRSVERVDAREEM- 307
DB 256 ELALMKLPK-SEALSARMLLIGAGKMGKLVIKHLVAKGCKVNVNRSVERVDAREEM- 315
QY 308 KDIEIVRPLSDMYQAAAEADVFTSTASETSLFAKEHAELPPVSDTMGGVRLFPVDISV 367
DB 316 PGVEIYKPLDEMLSCAAEADVFTSTASETSLFAKEHAELPPVSDTMGGVRLFPVDISV 371
QY 368 PRNVSCVSEVGAARVYVVDLKEVVEANKEDRLRKAMEAQTIIIEELRRFEAWRDSLET 427
DB 372 PRNVSCVSEVGAARVYVVDLKEVVEANKEDRLRKAMEAQTIIIEELRRFEAWRDSLET 431
QY 428 VPTIKLRSYADRIRASELEKLOKVGEDALTKQMRRAIEELSTGIVNKLHGPLOHLRC 487
DB 432 VPTIKLRSYADRIRASELEKLOKVGEDALTKQMRRAIEELSTGIVNKLHGPLOHLRC 490

RESULT 8
Ti0245
glutamyl-tRNA reductase (EC 1.2.1.-) 2 - cucumber
C:Species: Cucumis sativus (cucumber)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: Ti0245
R:Tanaka, R.; Yoshida, K.; Nakayashiki, T.; Masuda, T.; Taji, H.; Inokuchi, H.; Tanai,
Plant Physiol. 110, 1223-1230, 1996
A:Title: Differential expression of two hema mRNAs encoding glutamyl-tRNA reductase p
A:Reference number: Z16983; MUID:97088717; PMID:8934625
A:Accession: Ti0245
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-542 <TAN>
A:Cross-references: EMBL:D67088; NID:G1015318; PIDN:BAAL1091.1; PID:G1015319
A:Experimental source: cv. Aonagajibal; greening cotyledons
C:Genetics:
A:Gene: hema2
A:Genome: nuclear
C:Function:
A:Description: catalyzes the reduction of glutamyl-tRNA(Glu) by NADPH to glutamic ac
A:Pathway: aminolevulinate biosynthesis; porphyrin biosynthesis
C:Superfamily: glutamyl-tRNA reductase
C:Keywords: aminolevulinate biosynthesis; chloroplast; NADP; oxidoreductase; porphyrin

Query Match 60.4%; Score 1573; DB 2; Length 542;
Best Local Similarity 66.6%; Pred. No. 1.7e-80;
Matches 313; Conservative 74; Mismatches 79; Indels 4; Gaps 4;

217 ISTLFSEKANKAGKRVRAQNTIASGAVSVSSAAVELALTKLPGSVS--SAAHMLVIGAGEMGK 275F

277 LVIKHLVAKGCKVWVNVRSVERVDATIREEMKD-IETVVRPLSDMYQAAAEADVFTSTA 335F

276 RIIEHLVAKGCTKWVNVRSDEKVAIRKENQSGVEIYKPLDEILACAEANVFTSTS 335F

336 SETSLFAKEHAELPPVSDTNGVRLFYDISVPRNVSAACVSEVGAARVYVNVDDLKEVVEA 395F

336 SETPLFLKEHVEIILPPCPADY--ARLFVDISVPRNVSGVAEALDSARVYVNVDDLKEVVA 393F

396 NKEDRLKAMEAQIITTEELRRFPAWRSDLETVPPIKKLRSVADRIRASELSKCLKOVGE 455F

394 NKEDRARKSEALPIIREETIEPFGWRDSLQFTPIRKLRSKTERIRAECEVKLISKHG- 452F

456 DALTKORRAIEELSTGVNKLHGLHLCDCGSDSRTLDETLENMHALNMFSLDMEK 515F

453 NGMDKKTREAVEKQRIIVNNILDYPMKHLRYDGTGSSKLTLENMQAVNRVIELDGE- 511F

516 AITEQIKIAKVEK 528

512 -LLEEKIREKDK 523

RESULT 10

S77180

glutamyl-tRNA reductase (EC 1.2.1.1-) - *Synechocystis* sp. (strain PCC 6803)

N;Alternate names: hema protein; protein slr1808; transfer RNA-Gln reductase

C;Species: *Synechocystis* sp.

A;Variety: PCC 6803

C;Date: 25-Apr-1997 #sequence revision 25-Apr-1997 #text_change 16-Jun-2000

C;Accession: S77180; A38087; S37660

R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyao, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, DNA Res. 3, 109-136, 1996

A;Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp.

A;Reference number: S74322; MUID: 97061201; PMID: 8905231

A;Accession: S77180

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-463 <KAN>

A;Cross-references: EMBL:D90908; GB:AB001339; NID:gl652725; PIDN:BAAI7738.1; P

A;Note: protein slr1808

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1

R;Verkemp, E.; Jahn, M.; Jahn, D.; Kumar, A.M.; Soell, D.

J. Biol. Chem. 267, 8275-8280, 1992

A;Title: Glutamyl-tRNA reductase from *Escherichia coli* and *Synechocystis* 6803.

A;Reference number: A38087; MUID: 92235044; PMID: 1569081

A;Accession: A38087

A;Molecule type: DNA

A;Residues: 37-463 <R>

A;Cross-references: GB:M84218; NID:gl54469; PIDN:AAA27289.1; PID:gl54470

R;Grimm, B.

Vecedita 117 195-107 1902

A;Title: Identification of a hemaA gene from *Synechocystis* by complementation
A;Reference number: S37660; MUID:9303989; PMID:1459859
A;Accession: S37660
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 37-463 <GRI>
A;Cross-references: EMBL:X65963; NID:g288421; PIDN:CAA46779.1; PID:g288422
C;Genetics:
A;Gene: hemaA
A;Start codon: GTG
C;Function:
A;Description: catalyzes the reduction of glutamyl-tRNA(Glu) by NADPH to glutamate
A;Pathway: aminolevulinate biosynthesis; porphyrin biosynthesis
C;Superfamily: glutamyl-tRNA reductase
C;Keywords: aminolevulinate biosynthesis; NADP; oxidoreductase; porphyrin biosynthesis

Query Match 34.9%; Score 910.5; DB 1; Length 463;
Best Local Similarity 43.0%; Pred. No. 1.3e-43;
Matches 200; Conservative 85; Mismatches 163; Indels 17; Gaps 5.

[illegible]

Db 181 RTANITILGAGKMSRLLVQHLLSKRVKIDINIVRSVERAKLLVDQFKEANINIYNLSLK 240

Qy 322 QAAAEADVFTSTASSETSLFAKEHAELPPVSDTGMGVRFLFDVDSVPRNVSAACVSEVCA 381

Db 241 TILQNSDLIVFTGSSQSEPIITPE-----LINCNDLPSBELMLFDIAVPRNVPNVSPQONI 296

Qy 382 RVYNVDLKEVVEANKEDRLKAMEAQIITTEELRFEAWRDSLETPTIKKLRYADRI 441

Db 297 KVFNVDDLKVVSQNOQTRKMAKAAEILLEEELSAFNWGSLEAITINKLREKARI 356

Qy 442 RASELEKCLKVGEDALTKMRAIEELSTGIYNKLLHGPQLHRCDSGDSRTLDLEN 501

Db 357 RVKELEKAIISLGNFVSDH-QEIVESLTRGIVNKILHDPVQLRA-----QODIEIRGA 411

Qy 502 MHALNMFSLDMEK 515

Db 412 LKILQTLFNLTIK 425

RESULT 13

T31441

probable glutamyl-tRNA reductase (EC 1.2.1.1-) hema - Helicobacillus mobilis

C:Species: Helicobacillus mobilis

C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 15-Sep-2000

C:Accession: T31441

R:Xiong, J.; Inoue, K.; Bauer, C.E.

Proc. Natl. Acad. Sci. U.S.A. 95, 14851-14856, 1998

A:Title: Tracking molecular evolution of photosynthesis by characterization of a major P

A:Reference number: Z21036; MUID:99061957; PMID:9843979

C:Accession: T31441

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-443 <XIO>

A:Cross-references: EMBL:AF080002; NID:G3820336; PID:G3820540; PIDN:AAC04013.1

C:Gene: hema

C:Superfamily: glutamyl-tRNA reductase

C:Keywords: oxidoreductase

Query Match 24.7%; Score 643.5; DB 2; Length 443;
Best Local Similarity 34.4%; Pred. No. 9.6e-29;
Matches 157; Conservative 98; Mismatches 177; Indels 25; Gaps 9;

Qy 82 IAVIGLSVHTAPVEMREKLAELWPAICELTSLNHIIEAAVLSTCNRMIEIYVVALSW 141

Db 3 IFVVLNKHSAPEVREKLSFTEAQLSEALHKLQMGAGIECCILSTCNRTIELYQASTDM 62

Qy 142 NRGIREVVDWMKSGIPASERLHFLIRSSDATRHLFEVSAGLDSLVLGEGQILAQVK 201

Db 63 EKGMTAVKRFVLEWQLOPQDFSKFYVHTLYDAIRHLFRVASGLDSMVLGETQILGQVR 122

Qy 202 QVVRSGQSGGLGNIDRMFKDAITAGKRVSETNISGAVSVSAAVELALMKLPKSEA 261

Db 123 TAYQRSCNEDCSNGIVNTWFOQATVGRKVRTEGIDQHPVSIYSYAVELAEQVLGLGK 182

Qy 262 LSARMLLIGAGKMGKLVIKHLVAKGCKVNVNRSVERVDAIREMKDIEIYVRLPSDMY 321

Db 183 RTA--MVLGACKMSVLTIKHLVAGVDKIIIANRSVEKAEELAKSCGGEALSF---ADVN 237

Qy 322 QAAAEADVFTSTASSETSLFAKEHAELPPVSDTGMGVRFLFDVDSVPRNVSAACVSEVGA 380

Db 238 HRLEADILISCTAATHYVIRKSMVE---QVMDRRGKPKVPFFDIAVPRDIDPEVAQVPG 294

Qy 381 ARVNVVDLKEVVEANKEDRLKAMEAQIITTEELRFEAWRDSLETPTIKKLRYADRI 440

Db 295 THLYDIDAMQVIVDRNLAEKCAEAEIIEHEINQFMWINSFLVPIPTIVGLKNKGQ 354

Qy 441 IRASELEKCLKVGEDALTKMRAIEELSTGIYNKLLHGPLOHLR--CDGSDSRTLDLET 498

Db 355 IKEKELDRALCKLKH--LSEKEKLVGLASSIYNQLLHDPITQLRHYAASPEGHLYSEI 412

Qy 499 LENMHALNMFSLDM-----EKAIIEQKIKAKVEKTON 531

Db 413 LQN-----LFCLDVPGQRKHVVVH--YPAVEQRQN 441

RESULT 14

CS1183

glutamyl-tRNA reductase NMB0576 [imported] - Neisseria meningitidis (strain MC58 sero

C:Species: Neisseria meningitidis

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C:Accession: CS1183

R:Pettersen, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.

Hi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scariato, V.; Massignani, V.; Pizza, M.

Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;

A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: A81000; MUID:20175755; PMID:10710307

A:Accession: CS1183

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-415 <TET>

A:Cross-references: GB:A8002413; GB:A8002098; NID:G7225797; PIDN:AAF41004.1; PID:G722;

A:Experimental source: serogroup B, strain MC58

C:Genetics:

A:Gene: NMB0576

C:Superfamily: glutamyl-tRNA reductase

Query Match 24.5%; Score 637.5; DB 2; Length 415;
Best Local Similarity 34.3%; Pred. No. 1.9e-28;
Matches 149; Conservative 96; Mismatches 163; Indels 27; Gaps 9;

Qy 82 IAVIGLSVHTAPVEMREKLAELWPAICELTSLNHIIEAAVLSTCNRMIEIYVVALSW 141

Db 3 LTAVGLNHTQAPLSIREKLAPAAALPAVNRALRSNAATEAVILSTCNRTELYCVGDS- 61

Qy 142 NRGIREVVDWMKSGIPASERLHFLIRSSDATRHLFEVSAGLDSMVLGEGQILAQVK 201

Db 62 ----EEIIRWLADYHSLPIEIRPYLYALDMQETVRHAFRVACGLDSMVLGEPQILGQIK 117

Qy 202 QVVRSGQSGGLGNIDRMFKDAITAGKRVSETNISGAVSVSAAVELALMKLPKSEA 261

Db 118 DAVRAEQESMGKLNALFOKTFEVAKEVRTDTAVGNSVSMASASVKLAIEQIPDIDG 177

Qy 262 LSARMLLIGAGKMGKLVIKHLVAKGCKVNVNRSVERVDAIREMKD-IEIVYRP--LS 318

Db 178 LN--VLFIGAGEMIELVATYPAKSPRLMTVANRTLARA----QELCDKLGVAEPCLLS 231

Qy 319 DMYQAAAEADVFTSTASSETSLFAKEHAELPPVSDTGMGVRFLFDVDSVPRNVSAACVSE 377

Db 232 DLPAILHDYDVVVSSTASQLPIVCKGMVERALKQRQSM--PLFMLDLAVPRDIEAEVGD 288

Qy 378 VGARVNVVDLKEVVEANKEDRLKAMEAQIITTEELRFEAWRDSLETPTIKKLRSY 437

Db 289 LNDAYLTVDDMNVIVQSGKEAROKAAAAAETLVSEKVAEFVROOQGRQSVPLIKALRDE 348

Qy 438 ADRIRASELEKCLKVGEDALTKMRAIEELSTGIYNKLLHGPLOHLRCDGSDSRTLDE 497

Db 349 GEKARKQVLENAMKQLAGATABEV---LERLSVOLTNKLLHSPTQTLNKAGEEDKOL-- 403

Qy 498 TLENMHALNMFSLD 512

Db 404 ----VHAQAQIYHLD 414

RESULT 15

A81920

glutamyl-tRNA reductase (EC 1.2.1.1-) NMA0760 [imported] - Neisseria meningitidis (str

C:Species: Neisseria meningitidis

C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C:Accession: A81920

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo;

i; Holroyd, S.; Jägel, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandrea;

Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491.

Search completed: December 15, 2003, 12:56:40
Job time : 22 secs